

UNCLASSIFIED

AD-A265 252

2

SECURITY CLASSIFICATION OF THIS PAGE

1. REPORT SECURITY CLASSIFICATION DTIC ELECTE		2. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.	
3. SECURITY CLASSIFICATION UNCLASSIFIED		4. DECLASSIFICATION/DOWNGRADING SCHEDULE JUN1 1993	
5. PERFORMING ORGANIZATION REPORT NUMBER		6. MONITORING ORGANIZATION REPORT NUMBER	
7. NAME OF PERFORMING ORGANIZATION BOARD OF TRUSTEES UNIVERSITY OF ARKANSAS		8. NAME OF MONITORING ORGANIZATION DEPARTMENT OF DEFENSE AIR FORCE OFFICE OF SCIENTIFIC RESEARCH	
9. ADDRESS (City, State and ZIP Code) 120 Ozark Hall Fayetteville, AR 72701		10. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB, DC 20332-6448	
11. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR/NM		12. OFFICE SYMBOL (if applicable)	
13. ADDRESS (City, State and ZIP Code) Building 410 Bolling AFB, DC 20332-6448		14. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-89-0472	
15. TITLE (Include Security Classification) Computational Algorithms for Identification of Distributed Parameter Systems		16. SOURCE OF FUNDING NO.	
17. PERSONAL AUTHOR(S) Dennis W. Brewer and Robert K. Powers		18. PROGRAM ELEMENT NO. 6-111-15	
19. TYPE OF REPORT Final Technical		20. PROJECT NO. 3484	
21. TIME COVERED FROM 89/09/01 TO 93/02/28		22. TASK NO. D7	
23. DATE OF REPORT (Yr., Mo., Day) 93/04/24		24. WORK UNIT NO.	
25. TYPE OF REPORT Final Technical		26. PAGE COUNT 77	
27. SUPPLEMENTARY NOTATION			
28. COSATI CODES			
29. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) parameter estimation, fractional derivative damping			
30. ABSTRACT (Continue on reverse if necessary and identify by block number) This research established a general framework for the convergence of a parameter estimation algorithm based on quasilinearization which applies to a class of distributed parameter systems described by linear dynamical systems. Conditions were established which guarantee local convergence of the identification algorithm. The algorithm was applied to delay and coefficient identification in systems of delay-differential equations. Such systems have been proposed as hereditary models of aeroelastic systems. A numerical identification algorithm was developed and tested for estimating parameters in a Volterra integral equation arising from a viscoelastic model of a flexible structure with Boltzmann damping. In particular, one of the parameters identified was the order of the derivative in Volterra integro-differential equations containing fractional derivatives, a form of viscoelastic damping. A Galerkin approximation in the space variable was used to approximate the partial differential equation with memory by a system of integro-differential equations. Numerical experiments were performed to test the ability of the algorithm to estimate unknown damping parameters in these systems.			
31. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS <input type="checkbox"/>			
32. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED			
33. NAME OF RESPONSIBLE INDIVIDUAL DR MARC JACOBS		34. TELEPHONE NUMBER (include Area Code) (202) 767-5027	
35. OFFICE SYMBOL N/A			

DD FORM 1473, 83 APR

EDITION OF 1 JAN 73 IS OBSOLETE.

TION OF THIS PAGE

93-12234



UNCLASSIFIED

**COMPUTATIONAL ALGORITHMS FOR IDENTIFICATION OF
DISTRIBUTED PARAMETER SYSTEMS**

Dennis W. Brewer and Robert K. Powers
Department of Mathematical Sciences
University of Arkansas
Fayetteville, Arkansas 72701

24 April 1993

Final Technical Report
For Period 1 September 1989 - 28 February 1993

Prepared for

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Building 410
Bolling AFB, D.C. 20332

DTIC QUALITY INSPECTED 3,

Accession For	
NTIS CRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special

A-1

FINAL TECHNICAL REPORT

Introduction

The principal goal of this research effort was to develop an effective computational algorithm for the estimation of parameters in distributed parameter systems. The algorithm was developed in a general setting which allowed application to phenomena modeled by delay-differential equations, Volterra integral equations, and partial differential equations with memory terms.

In particular we investigated a gradient-based parameter estimation method for dynamical systems in an abstract space. The basic functional analytic framework was the theory of semigroups of operators in infinite dimensional space. This framework allowed application to distributed parameter systems modeled by hereditary systems and partial differential equations. This research focused both on theoretical aspects, such as convergence criteria, and on the efficient implementation and testing of the algorithms for computational purposes.

Summary of research

The dynamical systems under consideration were of the general form

$$\begin{aligned}\dot{x}(t) &= A(q)x(t) + Bu(t) \\ x(0) &= x_0 \\ y(t) &= Cx(t)\end{aligned}\tag{1}$$

where u and y are input and output functions, x is an infinite dimensional state, and $A(q)$ is an evolution operator depending on a possibly distributed parameter q . A computationally feasible algorithm was developed for solving the following identification problem.

Problem (ID). Given an input function u and observations y_i at times t_i , $i = 1, \dots, m$, find a system parameter q which minimizes the quadratic cost function

$$J(q) = \sum_{i=1}^m \|Cx(t_i; q) - y_i\|^2$$

where $x(t; q)$ is the state at time t of the dynamical system with parameter q .

Hereditary models of fluid-structure interaction often contain unknown parameters which need to be identified. Such models may be cast as a dynamical system in an infinite dimensional space as in equation (1). Our goal was to identify delays in these models as well as other system parameters. These results were developed jointly with J. A. Burns and E. M. Cliff of Virginia Tech University and published in a paper entitled, "Parameter identification for an abstract Cauchy problem by quasilinearization." For immediate reference, this paper is included in the appendix of this report.

Similarly, certain distributed parameter models of viscoelastic structures may also be formulated as an abstract dynamical system. The models of interest contain Boltzmann damping and, in particular, fractional derivative damping. The numerical phase of this research effort was principally concerned with the following aspects of the general setting described above:

- (1) the development and testing of a numerical algorithm for estimating parameters in a Volterra integral equation arising from a viscoelastic model of a flexible structure with Boltzmann damping;
- (2) the implementation of numerical methods for a system of Volterra equations resulting from a Galerkin approximation of a partial differential equation with hereditary effects.

Our principal research effort was directed toward the development and estimation of distributed parameter models of flexible structures with internal damping. The design of control systems for flexible structures is highly dependent on the amount of internal damping present in the structure. Damping parameters typically change as materials and geometries of the structures change. Accurate and efficient identification algorithms are needed to estimate a system's characteristics and implement a stable control algorithm.

In particular, we considered the partial differential equation

$$\rho u_{tt}(x,t) = \frac{\partial}{\partial x} \left\{ E u_x(x,t) + \frac{\partial}{\partial t} \int_0^t g(t-s) u_x(x,s) ds \right\} + f(x,t) \quad (2)$$

on $0 < x < 1$, $t > 0$, with appropriate boundary and initial conditions. The function $u(x,t)$ represents the longitudinal displacement at position x and time t along a uniform bar of density ρ where E is a stiffness parameter and $f(x,t)$ is a

forcing function. The function $g(s)$ models damping effects. The integral represents a Boltzmann-type internal damping term which assumes that the stress is function of the strain and the strain history. We studied fractional derivative damping models in which the kernel function is given by an expression of the form

$$g(s) = \frac{\gamma e^{-\beta s}}{\Gamma(1-\alpha) s^\alpha}, \quad s > 0, \beta > 0, 0 < \alpha < 1.$$

An estimation algorithm for a discretized form of equation (2) was formulated and tested. Using simulated data, it was shown to be means of identifying the parameters α and β .

A Galerkin approximation of equation (2) using, for example, linear splines in the space variable yields a system of Volterra integro-differential equations with an integrable, but unbounded, kernel. A gradient-based identification algorithm for the parameters α and β was implemented in a Volterra equation of the form

$$\begin{cases} \dot{w}(t) = Mw(t) + \int_{-\infty}^0 K(-s)w(t+s)ds + F(t), & t > 0, \\ w(0) = \eta, \quad w(s) = \varphi(s), & s < 0, \end{cases}$$

where $\eta = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}$, $M = \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}$, $F(t) = \begin{pmatrix} f(t) \\ 0 \end{pmatrix}$, and

$$K(s) = \frac{\gamma e^{-\beta s}}{\Gamma(1-\alpha) s^\alpha} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad s > 0.$$

This involved the implementation of numerical methods for solving the Volterra equation and its sensitivity equations with respect to the unknown parameters. Where possible, numerical results were checked against a closed-form solution obtained by a Laplace transform method using software capable of symbolic computation.

A gradient-based algorithm for identifying a singularity in a weakly-singular Volterra integral equation was established and numerically tested. These results were published in the *Journal of Integral Equations and Applications* in a paper referenced below. Additional numerical results were published in *Applied Numerical Mathematics*, also referenced below. For the immediate reference this paper is included in the appendix together with several figures which did not appear in the published version.

The Galerkin approximation of the hereditary partial

differential equation model in the space variable gives rise to a large system of Volterra equations with weakly singular kernels. Our experience indicated the importance of an accurate and efficient method for solving systems of this type with non-smooth solutions. We employed a product integration method in the time variable.

A class of fractional linear multi-step methods for the numerical solution of weakly singular Volterra equations was investigated, but not found to be significantly superior to product methods in this context. This was attributed to the fact that linear multi-step methods require a priori knowledge of the singularity. Numerical results indicated that the convergence rates of these methods deteriorate in the absence of this knowledge. In our context the singularity is among the parameters to be identified so is not known a priori. Our problem required a numerical integration scheme which is robust over a wide range of singularities.

The final numerical results of this research effort concerned a method for parameter estimation in a semi-discrete approximation of the partial differential equation

$$\left\{ \begin{aligned} \rho u_{tt}(x,t) &= E u_{xx}(x,t) \\ &+ \frac{\gamma}{\Gamma(1-\alpha)} \frac{\partial}{\partial t} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} u_{xx}(x,s) ds + f_1(x,t) \end{aligned} \right.$$

with boundary conditions $u(0,t) = 0$, $u(1,t) = 0$, and initial conditions $u(x,0) = u_0(x)$, $u_t(x,0) = u_1(x)$.

Integrating with respect to t one obtains

$$\left\{ \begin{aligned} \rho u_t(x,t) &= E \int_0^t u_{xx}(x,s) ds + \\ &\frac{\gamma}{\Gamma(1-\alpha)} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} u_{xx}(x,s) ds + f_2(x,t) \end{aligned} \right.$$

where $f_2(x,t) = \int_0^t f_1(x,s) ds + \rho u_1(x)$.

We applied a Galerkin approximation in which the interval $[0,1]$ is divided into N equal parts and the homogeneous boundary conditions allowed to approximate the solution by a function of the form $u(x,t) = \sum_{j=0}^N a_j(t) \phi_j(x)$ where ϕ_j is a cubic spline basis element. Substituting in the equation, taking an

inner product with a basis element ϕ_k , and integrating by parts yielded a Volterra equation in t of the form

$$\left\{ \begin{aligned} \rho A^N \dot{v}(t) &= -E D^N \int_0^t v(s) ds - \\ &\quad \frac{\gamma}{\Gamma(1-\alpha)} D^N \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} v(s) ds + f_3(t) \end{aligned} \right.$$

where A^N and D^N are $N+1 \times N+1$ matrices depending on inner products of ϕ_j and ϕ'_j .

The quasilinearization algorithm required that we solve this equation along with its sensitivity equations obtained by differentiating with respect to α and β , the unknown parameters. The sensitivity equations also have weakly singular kernels so that the same numerical methods may be applied. These results will be reported in the *Proceedings of the World Congress of Nonlinear Analysts*. This paper is included for reference in the appendix of this report.

Research articles

The following papers relating to this research effort were published or will soon appear in refereed journals or as invited papers in conference proceedings.

D. W. Brewer, Gradient methods for identification of distributed parameter systems, *Proceedings of the 28th IEEE Conference on Decision and Control*, December 1989, 599-603.

D. W. Brewer and R. K. Powers, Parameter identification in a Volterra equation with weakly singular kernel, *Journal of Integral Equations and Applications* 2(1990), 353-373.

D. W. Brewer and R. K. Powers, Parameter estimation for a Volterra integro-differential equation, *Applied Numerical Mathematics* 9(1992), 307-320.

D. W. Brewer and R. K. Powers, A direct method for parameter estimation in distributed systems, *Proceedings of the World Congress of Nonlinear Analysts*, August 1992, to appear.

D. W. Brewer, J. A. Burns, and E. M. Cliff, Parameter estimation for an abstract Cauchy problem by quasilinearization, *Quarterly of Applied Mathematics* 51(1993), 1-22.

Professional Personnel

The following persons were supported by this research effort.

D. W. Brewer, Co-Principal Investigator
R. K. Powers, Co-Principal Investigator
H. W. Lee, Graduate Research Assistant, Ph.D., August 1992
Thesis title: "Properties of solutions of a class of
Volterra and functional differential equations"

Interactions

The following papers relating to this research effort were presented at national and international meetings or regional conferences.

R. K. Powers, Identification of parameters in an Abel-type integro-differential equation, Midwest Conference on Differential Equations, Carbondale, Illinois, October 1989.

D. W. Brewer, Gradient methods for identification of distributed parameter systems, IEEE Conference on Decision and Control, Tampa, Florida, December 1989.

D. W. Brewer, Parameter identification for a Volterra integro-differential equation by quasilinearization, International Conference on the Numerical Solution of Volterra and Delay Equations, Tempe, Arizona, May 1990.

R. K. Powers, Identification of parameters in a weakly singular kernel arising in viscoelasticity, SIAM Annual Meeting, Chicago, Illinois, July 1990.

R. K. Powers, Identification of parameters for a Volterra integro-differential equation with weakly singular kernel, Memphis State - Arkansas Joint Conference in Analysis, Memphis, Tennessee, October 1990.

D. W. Brewer and R. K. Powers, Gradient methods for identification of a distributed parameter system, Second International Conference on Industrial and Applied Mathematics, Washington, D.C., July 1991.

R. K. Powers and D. W. Brewer, Numerical solution of a Volterra integro-differential equation arising in parameter identification, Second International Conference on Industrial and Applied Mathematics, Washington, D.C., July 1991.

D. W. Brewer, A direct method for parameter estimation in distributed systems, World Congress of Nonlinear Analysts, Tampa, Florida, August 1992.

In addition, D. W. Brewer made one-week research visits to the Interdisciplinary Center for Applied Mathematics (ICAM), Blacksburg, Virginia during the Spring of 1990, 1991, and 1992.

Appendix

Parameter estimation for an abstract Cauchy problem by quasilinearization, *Quarterly of Applied Mathematics* 51(1993), 1-22.

D. W. Brewer, J. A. Burns, and E. M. Cliff 9

Parameter identification in a Volterra equation with weakly singular kernel, *Journal of Integral Equations and Applications* 2(1990), 307-320.

D. W. Brewer and R. K. Powers 43

A direct method for parameter estimation in distributed systems, *Proceedings of the World Congress of Nonlinear Analysts*, to appear.

D. W. Brewer and R. K. Powers 67

PARAMETER IDENTIFICATION FOR AN ABSTRACT CAUCHY PROBLEM
BY QUASILINEARIZATION

Dennis W. Brewer

University of Arkansas, Fayetteville

John A. Burns

Virginia Polytechnic Institute and State University

Eugene M. Cliff

Virginia Polytechnic Institute and State University

Abstract. A parameter identification problem is considered in the context of a linear abstract Cauchy problem with a parameter-dependent evolution operator. Conditions are investigated under which the gradient of the state with respect to a parameter possesses smoothness properties which lead to local convergence of an estimation algorithm based on quasilinearization. Numerical results are presented concerning estimation of unknown parameters in delay-differential equations.

The research of the first and second authors was supported in part by the National Aeronautics and Space Administration under contract NAS1-18107 while the authors were in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23665. The research of the first author was supported in part by the Air Force Office of Scientific Research under grant AFOSR-89-0472. Additional support was received from DARPA under grant F49620-87-C-0116 while the first author was in residence at the Interdisciplinary Center for Applied Mathematics, Blacksburg, VA 24061. Also, research of the second and third authors was supported in part by AFOSR under grant AFOSR-89-0001, DARPA under grant F49620-87-C-0116, and SDIO under contract F49620-87-C-0088.

1. Introduction

During the past fifteen years considerable effort has been devoted to the problem of estimating unknown parameters in distributed parameter systems. The recent book by Banks and Kunisch [9] provides an excellent account of the progress made in the field. Many parameter estimation problems are best formulated as optimization problems (sometimes over infinite dimensional "parameter spaces") and algorithms are developed to minimize an appropriate cost function. Although there are several approaches to these problems, their infinite dimensional nature requires that numerical approximations be introduced at some point in the analysis. Consequently, there are two basic classes of algorithms for optimization based parameter estimation. The first type of algorithm, and the most frequently used for dynamic problems, is indirect and proceeds by initially approximating the dynamic equations (e.g. finite elements, finite differences, etc.) and then using optimization algorithms on the finite dimensional problem. This approach is typified by the papers [1]-[6], [8], [10], and [18].

The second more direct approach is based on the direct application of an optimization algorithm and employing numerical approximations at each step of the algorithm to compute the necessary solutions of the dynamic equations. This approach is used in [12], [13], [17], and [19]. Both methods have advantages and disadvantages. Depending on the particular type of distributed parameter system, one method may out perform the other.

Although we shall consider only the problem of identifying a finite number of parameters, the infinite dimensional dynamic constraint enters into the optimization algorithm. Basically, the objective function from parameter space to \mathbb{R} is a composition of a finite rank map with an operator (defining the dynamic constraint) on an infinite dimensional space. Therefore, any method that requires gradients to be computed will have to deal with the differentiation of the infinite dimensional constraint, *i.e.* the chain rule is needed. It is in this sense that the quasilinearization algorithm considered here has an "infinite dimensional" nature.

Direct methods such as quasilinearization considered here are often limited by the fact that the dependence on unknown parameters of the solution to the infinite dimensional dynamical equations may not be "smooth enough" to establish convergence of the algorithm. Indeed, some algorithms may not be properly defined without this necessary smoothness. Indirect methods avoid this difficulty and often lead to easily implemented algorithms. On the other hand, when direct methods can be applied it is sometimes possible to establish the convergence and the rates of convergence to the unknown optimal parameters (see [13], [19]).

This paper considers the dependence on an unknown parameter q of the solution of the linear abstract Cauchy problem

$$(1.1) \quad \begin{cases} \dot{x}(t) = A(q)x(t) + u(t), & 0 \leq t \leq T, \\ x(0) = x_0. \end{cases}$$

Our ultimate goal is to formulate and establish the convergence of a gradient-based parameter estimation algorithm applicable in this abstract setting.

This algorithm employs computation of the gradient $D_q x(t; q)$ of the solution of (1.1) with respect to the parameter. Conditions for the existence of this gradient are established in [11]. In Section 2 we review these conditions and the general setting for the remainder of the paper. Convergence of the algorithm requires certain smoothness properties of the gradient $D_q x(t; q)$ with respect to q . These properties are established in Section 3 and their applicability to a linear delay-differential equation is discussed in Section 4. In this example the delay is among the parameters so that in this setting the parameter dependence appears in unbounded terms of the evolution operator $A(q)$.

An abstract parameter estimation algorithm for a finite dimensional parameter space using a discrete cost function is presented in Section 5. In Section 6 its convergence is established using the results of Section 3. In Section 7 we present several numerical examples which indicate the performance of the algorithm for delay and coefficient estimation in linear delay-differential equations. Additional examples may be found in [12].

Numerical testing and evaluation on a wider variety of parameter estimation problems will be undertaken in a subsequent paper.

2. The General Setting

The application of quasilinearization to parameter estimation requires knowledge of the derivative of the state with respect to the unknown parameter. This topic is addressed in [11]. In this section we review the framework used there to obtain differentiability and establish notation to be used in the remainder of this paper.

Let P be an open subset of a normed linear space \mathbb{P} with norm $|\cdot|$ and let X be a Banach space with norm $\|\cdot\|$. For every $q \in P$ let $A(q)$ be a linear operator on $D(A(q))$ in X . Throughout this paper we assume

(H1) $A(q)$ generates a strongly continuous semigroup $S(t; q)$ on X ;

(H2) $D(A(q)) = D$ is independent of q ;

(H3) $\|S(t; q)x\| \leq Me^{\omega t}\|x\|$, $x \in X$, $t \geq 0$, $q \in D$, for some constants M and ω independent of q , x , and t .

Fix $T > 0$ and $u \in L^1(0, T; X)$. Define $Q(t; q) = \int_0^t S(t-s; q)u(s)ds$ for $q \in P$,

$0 \leq t \leq T$. Note that if (1.1) has a strong solution then it is given by the formula $x(t) = S(t; q)x_0 + Q(t; q)$ for $0 \leq t \leq T$.

In applications of this theory it is useful to consider just those terms of $A(q)$ in which the parameter appears. To this end we write $A(q) = A + B(q)$ where A and $B(q)$ both have domain D and A is independent of q . Concerning $B(q)$ we assume the following:

(H4) For every $q, q_0 \in P$ there is a constant K such that

$$\int_0^T \|B(q)S(t; q_0)x\| dt \leq K\|x\| \quad \text{for all } x \in D.$$

In Section 4 we discuss an example in which an unbounded operator $B(q)$ satisfies (H4). This hypothesis does imply, however, that the linear mapping $x \rightarrow B(q)S(\cdot; q_0)x$ is bounded as a mapping from D into $L^1(0, T; X)$. Let $F(q, q_0)$ denote the bounded linear extension of this operator to X . Let $\|\cdot\|_1$ denote the norm in $L^1(0, T; X)$. Concerning F we assume the following:

(H5) There is closed subspace Y of X such that

(i) $F(q, q_0)x_0 \in L^1(0, T; Y)$ for $q, q_0 \in P$, and

(ii) for every $q_0 \in P$ and $\epsilon > 0$ there exists $\delta > 0$ such that

$$\|F(q, q_0)y - F(q_0, q_0)y\|_1 \leq \epsilon \|y\| \text{ for } y \in Y \text{ and}$$

$$|q - q_0| \leq \delta.$$

The analogue of F for the function $Q(t; q)$ is the mapping $G(q, q_0)$ from $L^1(0, T; D)$ into $L^1(0, T; X)$ defined by

$$[G(q, q_0)w](t) = \int_0^t B(q)S(t-s; q_0)w(s)ds.$$

By (H4) it follows that G can be extended to a bounded linear mapping on $L^1(0, T; X)$ so that in particular $G(q, q_0)u$ is defined as an element of $L^1(0, T; X)$. In addition we assume

(H6) $G(q, q_0)u \in L^1(0, T; Y)$ for $q, q_0 \in P$

where Y denotes the subspace required by (H5).

3. Parameter Dependence

In this section we deduce smoothness properties of the solution $x(t;q) = S(t;q)x_0 + Q(t;q)$ with respect to q . These properties are derived from similar properties of $F(q,q_0)$ and $G(q,q_0)$ which are operators related to $A(q)$. These results will be used in Section 5 to prove convergence of the parameter estimation algorithm. Throughout this section $T > 0$, $x_0 \in X$, and $u \in L^1(0,T;X)$ are fixed as given in (1.1). The symbol D_q denotes Frechet differentiation with respect to q . These results are given as a series of lemmas whose proofs are at the end of this section.

Lemma 3.1. *Suppose (H1) – (H5) hold. In addition, suppose that for a given $q^* \in P$*

(H7) $F(q,q_0)x_0$ is Frechet differentiable with respect to q at q_0
for every $q_0 \in P$.

For brevity, let $DF(q_0)$ denote $D_q[F(q,q_0)x_0]|_{q=q_0}$ for $q_0 \in P$. In addition, suppose

(H8) $DF(q)$ is strongly continuous in q at q^* , that is, for each $h \in P$ the mapping $q \rightarrow DF(q)h$ from P into $L^1(0,T;X)$ is continuous at q^* .

Then for each $t \in [0,T]$, $S(t;q)x_0$ is Frechet differentiable with respect to q at every $q \in P$ and $D_q[S(t;q)x_0]$ is strongly continuous with respect to q at q^* .

Lemma 3.2. *Suppose (H1) – (H6) hold and in addition suppose that for a given $q^* \in P$,*

(H9) $G(q, q_0)u$ is Frechet differentiable with respect to q at q_0
for every $q_0 \in P$.

Again denoting this derivative by $DG(q_0)$ for $q_0 \in P$, assume

(H10) $DG(q)$ is strongly continuous in q at q^* .

Then for $t \in [0, T]$, $Q(t; q)$ is Frechet differentiable with respect to q at every $q \in P$ and $D_q[Q(t; q)]$ is strongly continuous in q at q^* .

Lemma 3.3. Suppose (H1) – (H5) and (H7) hold and in addition suppose

(H11) $F(q, q^*)$ is locally Lipschitz continuous in q at q^* , uniformly
for $y \in Y$, that is, there exist constants $K_1, \delta_1 > 0$ such that

$$\|F(q, q^*)y - F(q^*, q^*)y\|_1 \leq K_1 |q - q^*| \|y\|$$

whenever $|q - q^*| < \delta_1$ and $y \in Y$.

Moreover, assume that

(H12) $DF(q)$ is strongly locally Lipschitz continuous with respect
to q at q^* . That is, for each $h \in P$, there are constants
 $K, \delta > 0$ such that

$$\|DF(q)h - DF(q^*)h\| \leq K|q - q^*|$$

for $|q - q^*| \leq \delta$.

Then $D_q[S(t; q)x_0]$ is strongly locally Lipschitz continuous with respect to
 q at q^* for every $t \in [0, T]$.

Lemma 3.4. Suppose (H1) - (H6), (H9) - (H10) hold and in addition suppose

(H13) $DG(q)$ is strongly locally Lipschitz continuous with respect to q at q^* .

Then $D_q[Q(t;q)]$ is strongly locally Lipschitz continuous with respect to q at q^* for every $t \in [0, T]$.

Although the assumptions (H1) - (H13) are rather technical, we shall see that they can be easily verified for delay systems even in the case that the unknown parameter is the delay itself. Therefore, the results presented here remove the limitations placed on the perturbation $B(q)$ in papers [13] and [16].

For completeness we now present the proofs of Lemma 3.1 - Lemma 3.4. However, these proofs make use of the basic results found in [11] and in order to keep the length of the proofs reasonable we assume that the reader has [11] in hand.

Proof of Lemma 3.1. It is shown in [11] that (H1) - (H5), (H7) imply that $D_q[S(t;q)x_0]$ exists for $q \in P$. Furthermore, it is given by the formula

$$(3.1) \quad D_q[S(t;q)x_0]h = \int_0^t S(t-s;q)[DF(q)h](s)ds, \quad h \in P.$$

We therefore obtain by substitution

$$(3.2) \quad \begin{aligned} D_q[S(t;q)x_0]h - D_q[S(t;q^*)x_0]h \\ = \int_0^t [S(t-s;q) - S(t-s;q^*)][DF(q)h](s)ds \\ + \int_0^t S(t-s;q^*)([DF(q)h](s) - [DF(q^*)h](s))ds. \end{aligned}$$

Let $\epsilon > 0$ be given and let $C = Me^{\omega t}$. It can be shown (see the proof of Theorem 1 [11]) that for all $x \in X$

$$(3.3) \quad \|S(t; q)x - S(t; q^*)x\| \leq C\|F(q, q^*)x - F(q^*, q^*)x\|_1.$$

Combining (3.3) with (H5ii) shows that for some $\delta_1 > 0$

$$\|S(t, q)y - S(t; q^*)y\| \leq \epsilon C\|y\|, \quad 0 \leq t \leq T, \quad y \in Y,$$

whenever $|q - q^*| \leq \delta_1$. In particular, putting $y = [DF(q)h](s) \in Y$ by (H5i) we obtain

$$\|[S(t-s; q) - S(t-s; q^*)][DF(q)h](s)\| \leq \epsilon C\|[DF(q)h](s)\|$$

for $|q - q^*| \leq \delta_1$, a.e. $s \in (0, T)$. Since $DF(q)h$ is continuous at q^* , there exist constants K_2 , $\delta_2 > 0$ such that

$$\|DF(q)h\|_1 \leq K_2 \text{ for } |q - q^*| \leq \delta_2.$$

Combining these estimates shows that the first term in (3.2) is bounded by ϵCK_2 if $|q - q^*| \leq \min(\delta_1, \delta_2)$.

Using (H8) it is easy to see that there exists $\delta_3 > 0$ such that the second term in (3.2) is bounded by ϵC for $|q - q^*| \leq \delta_3$. These estimates complete the proof of Lemma 3.1.

Proof of Lemma 3.2. By Theorem 3 of [11], $D_q[Q(t;q)]$ exists for $q \in P$ and

$$(3.4) \quad D_q[Q(t;q)] - D_q[Q(t;q^*)] \\ = \int_0^t [S(t-s;q) - S(t-s;q^*)][DG(q)(s)]ds \\ + \int_0^t S(t-s;q^*)[(DG(q))(s) - (DG(q^*))(s)]ds$$

where u has been suppressed in the notation. Since $DG(q) \in L^1(0,T;Y)$ for $q \in P$ by (H6), the proof follows exactly as in the proof of Lemma 3.1.

Proof of Lemma 3.3. Let $\epsilon > 0$ be given. By (3.3) and (H11) there exists $\delta_1 > 0$ such that

$$\|S(t;q)y - S(t;q^*)y\| \leq CK_1 \|y\| |q - q^*|$$

for $y \in Y$ and $|q - q^*| \leq \delta_1$. Since $DF(q)h \in L^1(0,T;Y)$ by (H5i) we have as in the proof of Lemma 2.1 that the first term of (3.2) is bounded by $K_1 K_2 |q - q^*|$ for $|q - q^*| \leq \min(\delta_1, \delta_2)$. An estimate of the same form is easily obtained for the second term of (3.2) using (H12). These estimates complete the proof of Lemma 3.3.

Proof of Lemma 3.4. Since $DG(q)u \in L^1(0,T;Y)$ by (H6), the proof follows exactly as in the proof of Lemma 3.3 using (3.4) in place of (3.2).

4. Application to a Delay-Differential Equation

In this section we apply the framework of the previous sections to the linear delay-differential equation

$$(4.1) \quad \begin{cases} \dot{x}(t) = a_0 x(t) + \sum_{k=1}^n a_k x(t - q_k) + u(t) \\ x(0) = \eta \\ x_0 = \varphi. \end{cases}$$

Let $\mathbb{P} = \mathbb{R}^n$, fix $r > 0$, and let $P = \{q = (q_1, q_2, \dots, q_n) : 0 < q_k < r \text{ for } k = 1, 2, \dots, n\}$. In equation (4.1), $\eta \in \mathbb{R}$, $a_k \in \mathbb{R}$, $k = 0, 1, \dots, n$, $\varphi \in L^1(-r, 0)$ with norm denoted by $\|\varphi\|_1$, $u \in L^1(0, T)$, and $x_t(s) = x(t+s)$ for $t \geq 0$, $-r \leq s \leq 0$. By a solution of (4.1) we mean a function x which is absolutely continuous on $[0, T]$ and satisfies (4.1) almost everywhere on $(0, T)$.

Following the construction in [14], we take $X = \mathbb{R} \times L^1(-r, 0)$ with norm $\|(\eta, \varphi)\| = |\eta| + \|\varphi\|_1$ and define for $q \in P$ an operator $A(q)$ on

$$D = \{(\eta, \varphi) \in X : \varphi \text{ is abs. cont. on } [-r, 0], \dot{\varphi} \in L^1(-r, 0), \text{ and } \varphi(0) = \eta\}$$

by

$$A(q)(\eta, \varphi) = (a_0 \varphi(0) + \sum_{k=1}^n a_k \varphi(-q_k), \dot{\varphi}).$$

Then it is well known that $A(q)$ generates a strongly continuous semigroup $S(t; q)$ on X satisfying $S(t; q) = (y(t), y_t)$ where $y(t) = y(t; q)$ denotes the solution of (4.1) with $u = 0$. It is a consequence of standard results that (H1) - (H3) hold in this setting.

For $q = (q_1, \dots, q_n)$ and q_0 in P , $(\eta, \varphi) \in X$, and $w \in L^1(0, T)$ it follows that in this example the mappings F and G of Section 3 are given by

$$(4.2) \quad F(q, q_0)(\eta, \varphi) = \left(\sum_{k=1}^n a_k y(\cdot - q_k; q_0), 0 \right)$$

and

$$(4.3) \quad [G(q, q_0)w](t) = \left(\sum_{k=1}^n a_k z(t - q_k; q_0), 0 \right)$$

for a.e. $t \in (0, T)$ where $z(t; q)$ denotes the solution of (4.1) with $u = w$ and $(\eta, \varphi) = (0, 0)$. It is shown in [11] that these mappings satisfy (H4) - (H6) with the closed subspace $Y = \mathbb{R} \times \{0\}$. It is also shown in [11] that F and G satisfy the differentiability hypotheses (H7) and (H9) for $(\eta, \varphi) = x_0 \in D$ and $q, q_0 \in P$. Furthermore, their Frechet derivatives are given by

$$(4.4) \quad [DF(q)h](t) = \left(- \sum_{k=1}^n a_k \dot{y}(t - q_k; q) h_k, 0 \right)$$

and

$$(4.5) \quad [DG(q)h](t) = \left(- \sum_{k=1}^n a_k \dot{z}(t - q_k; q) h_k, 0 \right)$$

for $q \in P$, $h = (h_1, \dots, h_n) \in \mathbb{R}^n$, where $y(t; q)$ is the solution of (4.1) with $u = 0$ and $z(t; q)$ is the solution of (4.1) with $(\eta, \varphi) = (0, 0)$.

It remains to establish conditions under which (H8), (H10) - (H13) are satisfied.

Lemma 4.1. Fix $q^* = (q_1^*, \dots, q_n^*) \in P$ and $x_0 \in D$. Then $F(q, q^*)x_0$ as defined by (4.2) satisfies (H11).

Proof: In Section 5 of [11] it is shown that there are constants C_2 and

$\delta_2 > 0$ such that

$$\|F(q^*+h, q^*)(\eta, 0) - F(q^*, q^*)(\eta, 0)\|_1 \leq C_2 |h| \|(\eta, 0)\|$$

for $h \in \mathbb{R}^n$, $\eta \in \mathbb{R}$, $|h| \leq \delta_2$. Here we define $|h| = \sum_{k=1}^n |h_k|$. This estimate is equivalent to (H11) with $Y = \mathbb{R} \times \{0\}$.

Lemma 4.2. Suppose $x_0 = (\eta, \varphi) \in D$. Then $DF(q)$ as given by (4.4) satisfies

(H8). Moreover, if in addition $\dot{\varphi}$ is of bounded variation on $[-r, 0]$, then $DF(q)$ satisfies (H12).

Proof: Let $A_m = \max_k |a_k|$ and $|h| = \max_k |h_k|$. Then we obtain the estimate

$$(4.6) \quad \|DF(q)h - DF(q^*)h\|_1 \leq A_m |h| \sum_{k=1}^n \int_0^T |\dot{y}(t-q_k; q) - \dot{y}(t-q_k; q^*)| dt \\ + A_m |h| \sum_{k=1}^n \int_0^T |\dot{y}(t-q_k; q^*) - \dot{y}(t-q_k; q^*)| dt.$$

Now from (4.1) we obtain

$$(4.7) \quad \int_0^T |\dot{y}(t-q_k; q) - \dot{y}(t-q_k; q^*)| dt \leq \int_0^T |\dot{y}(t; q) - \dot{y}(t; q^*)| dt \\ \leq A_m \sum_{j=1}^n \int_0^T |y(t-q_j; q) - y(t-q_j^*; q^*)| dt \\ \leq A_m \sum_{j=1}^n \int_0^T |y(t-q_j; q) - y(t-q_j; q^*)| dt \\ + A_m \sum_{j=1}^n \int_0^T |y(t-q_j; q^*) - y(t-q_j^*; q^*)| dt$$

$$\begin{aligned} &\leq A_m \sum_{j=1}^n \int_0^T |y(t; q) - y(t; q^*)| dt \\ &\quad + A_m \sum_{j=1}^n \int_0^T |y(t - q_j; q^*) - y(t - q_j^*; q^*)| dt. \end{aligned}$$

Now since $y(t; q) = S(t; q)x_0$ is differentiable with respect to q it is not difficult to show that there are constants β and δ such that

$$(4.8) \quad \int_0^T |y(t; q) - y(t; q^*)| dt \leq \beta |q - q^*|$$

whenever $|q - q^*| \leq \delta$. Combining (4.7) and (4.8) with (4.6) yields

$$\begin{aligned} (4.9) \quad \|DF(q)h - DF(q^*)h\|_1 &\leq A_m^2 |h| n \beta |q - q^*| \\ &\quad + A_m^2 |h| n \sum_{k=1}^n \int_0^T |y(t - q_k; q^*) - y(t - q_k^*; q^*)| dt \\ &\quad + A_m |h| \sum_{k=1}^n \int_0^T |\dot{y}(t - q_k; q^*) - \dot{y}(t - q_k^*; q^*)| dt. \end{aligned}$$

Since $(\eta, \varphi) \in D$, we have y and \dot{y} in $L^1(-r, T)$. Therefore, the integral terms in (4.9) approach zero as $q \rightarrow q^*$ and (H8) holds. If $\dot{\varphi}$ is of bounded variation on $[-r, 0]$, then y and \dot{y} are of bounded variation on $[-r, T]$. By [15, Theorem 2.1.7(b)] this implies that the integral terms in (4.9) are $O(|q - q^*|)$ as $q \rightarrow q^*$ so that (H12) holds.

Lemma 4.3. Suppose $u \in L^1(0, T)$. Then $DG(q)$ as defined by (4.5) satisfies (H10). Moreover, if in addition u is of bounded variation on $[0, T]$, then $DG(q)$ satisfies (H13).

Proof: Using (4.5) in place of (4.4) one obtains the estimate (4.9) above with y replaced by z . Now if $u \in L^1(0,T)$ then z and \dot{z} are in $L^1(-r,T)$ so that (H10) holds. Similarly, if u is of bounded variation on $[0,T]$, then z and \dot{z} are of bounded variation on $[-r,T]$ so that (H13) is satisfied.

5. The Algorithm

In this section we define an estimation algorithm over a finite dimensional parameter space based on quasilinearization and establish local convergence using the results of Section 3. In particular, we assume that the parameter space P is \mathbb{R}^n with canonical basis e_i , $i = 1, 2, \dots, n$. This algorithm can also be cast in a separable Hilbert space as in [17].

Given $x_0 \in D$ and $q \in P \subset \mathbb{R}^n$ a strong solution of (1.1) is given by $S(t;q)x_0 + Q(t;q)$. Here we have used the notation of Section 2. Let C be a bounded linear mapping from X into \mathbb{R}^ℓ and define

$$\gamma(t;q) = C[S(t;q)x_0 + Q(t;q)].$$

The parameter estimation algorithm is related to the following optimization problem.

Problem 5.1. Let $\bar{y}_j \in \mathbb{R}^\ell$, $j = 1, 2, \dots, m$ be data values taken at times $t_j \in [0, T]$, $j = 1, 2, \dots, m$, respectively. For $q \in P$ define the quadratic cost function

$$J(q) = \sum_{j=1}^m |\gamma(t_j;q) - \bar{y}_j|^2.$$

Find $q^* \in P$ such that $J(q^*) \leq J(q)$ for all $q \in P$.

The quasilinearization method defines a recursive algorithm whose fixed point is a local solution of Problem 5.1. A more complete exposition is given in [7]. Given an initial guess $q_0 \in P$ define

$$q_{k+1} = f(q_k), \quad k = 0, 1, 2, 3, \dots$$

where

$$f(q) = q - [D(q)]^{-1}b(q)$$

$$D(q) = \sum_{j=1}^m M^T(t_j; q) M(t_j; q)$$

$$b(q) = \sum_{j=1}^m M^T(t_j; q) [\gamma(t_j; q) - \bar{y}_j]$$

and the matrix $M(t; q)$ has its i^{th} column $M^i(t; q)$ given by

$$M^i(t; q) = CD_q[S(t; q)x_0 + Q(t; q)]e_i, \quad i = 1, 2, 3, \dots, n.$$

Lemma 5.1. Suppose the hypotheses of Lemmas 3.1 and 3.2 are satisfied.

Then $M(t_j; q)$ is continuous in q at q^* .

Proof. This is a direct consequence of Lemmas 3.1 and 3.2 and the above definitions.

Lemma 5.2. Suppose the hypotheses of Lemmas 3.3 and 3.4 are satisfied.

Then there exist constants $\alpha, \delta > 0$ such that

$$|M(t_j; q) - M(t_j; q^*)| \leq \alpha |q - q^*|.$$

for $|q - q^*| \leq \delta, \quad j = 1, 2, \dots, m.$

Proof. This is a direct consequence of Lemmas 3.3 and 3.4 and the above definitions.

Note that although the smoothness results of the previous sections hold for an infinite dimensional parameter, the implementation of the solution of Problem 5.1 by this method is limited to finitely many parameters. In fact a simple rank argument is used in [17] to show that if the number of parameters, n , exceeds the number of data values, m , then the matrix $D(q)$ is singular. In [17] a pseudo-inverse is proposed as a means of solving the underdetermined problem.

We can now prove the following convergence results. These results are typical of quasilinearization methods and the proofs given here are in the same spirit as those in [7]. We obtain superlinear convergence when there is an exact fit to data (Theorem 5.1) and linear convergence in the presence of error (Theorem 5.2).

Theorem 5.1. *Suppose the hypotheses of Lemmas 3.1 and 3.2 are satisfied.*

Moreover, assume $[D(q^)]^{-1}$ exists, $f(q^*) = q^*$, and $J(q^*) = 0$. Then for every $\epsilon > 0$ there exists $\delta > 0$ such that*

$$|f(q) - f(q^*)| \leq \epsilon |q - q^*|$$

for $|q - q^| \leq \delta$. In particular, there is a neighborhood U of q^* such that $q_k \rightarrow q^*$ as $k \rightarrow \infty$ whenever $q_0 \in U$.*

Proof. Note that $f(q^*) = q^*$ implies that $b(q^*) = 0$, or

$$(5.1) \quad \sum_{j=1}^m y_j^T(t_j; q^*) [\gamma(t_j; q^*) - \bar{y}_j] = 0.$$

Therefore

$$\begin{aligned}
f(q) - f(q^*) &= D(q)^{-1} [D(q)(q - q^*) - b(q)] \\
&= D(q)^{-1} \left[\sum_{j=1}^m M^T(t_j; q) [M(t_j; q)(q - q^*) - (\gamma(t_j; q) - \bar{y}_j)] \right] \\
&= D(q)^{-1} \sum_{j=1}^m M^T(t_j; q) [M(t_j; q) - M(t_j; q^*)](q - q^*) \\
&\quad - D(q)^{-1} \sum_{j=1}^m M^T(t_j; q) [\gamma(t_j; q) - \gamma(t_j; q^*) - M(t_j; q^*)(q - q^*)] \\
&\quad - D(q)^{-1} \sum_{j=1}^m M^T(t_j; q) [\gamma(t_j; q^*) - \bar{y}_j].
\end{aligned}$$

Therefore, using (5.1) we have that

$$(5.2) \quad f(q) - f(q^*) =$$

$$\begin{aligned}
&D(q)^{-1} \sum_{j=1}^m M^T(t_j; q) [M(t_j; q) - M(t_j; q^*)](q - q^*) \\
&- D(q)^{-1} \sum_{j=1}^m M^T(t_j; q) [\gamma(t_j; q) - \gamma(t_j; q^*) - M(t_j; q^*)(q - q^*)] \\
&- D(q)^{-1} \sum_{j=1}^m [M^T(t_j; q) - M^T(t_j; q^*)] [\gamma(t_j; q^*) - \bar{y}_j].
\end{aligned}$$

Note that $D(q)^{-1}$ exists and is bounded in a neighborhood of q^* since

$D(q^*)^{-1}$ exists by assumption and $D(q)^{-1}$ is continuous at q^* by Lemma 5.1.

Let $\epsilon > 0$ be given. Using Lemma 5.1 it is easy to see that there exist constants $\beta_1, \delta_1 > 0$ such that the first term in (5.2) is bounded by $\epsilon\beta_1|q - q^*|$ for $|q - q^*| \leq \delta_1$. Furthermore, since $M(t_j; q^*)$ is the Frechet derivative of $\gamma(t_j; q)$ at q^* , one can show that there exist constants $\beta_2, \delta_2 > 0$ such that the second term of (5.2) is bounded by $\epsilon\beta_2|q - q^*|$ for $|q - q^*| \leq \delta_2$. Combining these estimates with (5.2) yields

$$(5.3) \quad |f(q) - f(q^*)| \leq$$

$$\epsilon\beta|q - q^*| + |D(p)^{-1}| \sum_{j=1}^m |M^T(t_j; q) - M^T(t_j; q^*)| |\gamma(t_j; q^*) - \bar{y}_j|,$$

for $|q - q^*| \leq \delta = \min(\delta_1, \delta_2)$ and $\beta = \beta_1 + \beta_2$. Since $J(q^*) = 0$, the last term in (5.3) is zero. This estimate yields the desired result.

The following theorem does not require an exact fit to data, but does place some technical restrictions on the behaviour of M near q^* . Note that if Lemmas 3.3 and 3.4 hold then there exists $\bar{\delta} > 0$ such that for $0 < \delta < \bar{\delta}$ there exists a constant $K(\delta)$ such that

$$\sum_{j=1}^m |M^T(t_j; q) - M^T(t_j; q^*)| \leq K(\delta)|q - q^*|$$

for $|q - q^*| \leq \delta$. Let $K^* = \limsup_{\delta \downarrow 0} K(\delta)$ and define

$$(5.4) \quad \lambda^* = K^* |D(q^*)^{-1}| \max |\gamma(t_j; q^*) - \bar{y}_j|.$$

Theorem 5.2. Suppose the hypotheses of Lemmas 3.3 and 3.4 are satisfied.

Moreover, assume $[D(q^*)]^{-1}$ exists and $f(q^*) = q^*$. Let λ^* be defined by (5.4) and assume $\lambda^* < 1$. Then there exists $\delta^* > 0$ such that

$$|f(q) - f(q^*)| \leq \lambda^* |q - q^*|$$

for $|q - q^*| \leq \delta^*$. In particular, $q_k \rightarrow q^*$ as $k \rightarrow \infty$ whenever

$$|q_0 - q^*| \leq \delta^*.$$

Proof. This estimate is a direct consequence of (5.3).

6. Numerical Examples

In this section we consider several examples in which the above algorithm was used to solve parameter estimation problems in delay-differential equations. In these examples the emphasis is on delay identification since in the abstract setting this represents an unbounded perturbation of the generator as noted in Section 4.

With the exception of Example 6.8, the various unknown parameters are estimated using data generated from closed-form expressions for the solution found by the "method of steps". The algorithm is implemented by an averaging scheme [2] which approximates the state equation and the associated sensitivity equations by a system of ordinary differential equations. This system is solved by a fourth-order Runge-Kutta routine.

In the one delay examples the averaging scheme is implemented with the delay interval $[-r, 0]$ divided into sixteen equal segments, except that Example 6.8 uses 64 equal segments. In the two delay examples the intervals $[-r_2, -r_1]$ and $[-r_1, 0]$ are divided into sixteen equal segments. All computations were done on a VAX 11/750 minicomputer or a SUN Microsystems at the Institute for Computer Applications in Science and Engineering (ICASE).

Example 6.1. This example illustrates the rapid convergence of the method for a single unknown parameter--the delay in the following equation--with an initial guess which is an order of magnitude greater than the "true value" of $r = 1.0$. The equation and the results of the iteration are given below.

$$\begin{cases} \dot{x}(t) = -2x(t) + 3x(t-r), & t \geq 0 \\ x(t) = t + 1, & t \leq 0 \end{cases}$$

iterate	r	error
0	10.000	34.056
1	1.299	0.955
2	0.946	0.175
3	0.989	0.115
4	0.987	0.115

The convergence of the states to ten data points on the interval $[0,2]$ is illustrated in Figure 1.

Example 6.2. The data is the same as for Example 6.1, however in this case the algorithm is asked to estimate the coefficients as well as the delay. The equation shows an insensitivity to the individual coefficients which leads to the inaccuracy in the converged estimates. In fact, because of errors introduced by the averaging scheme for computing the state, the estimated values fit the data better than the "true values" used to compute the data by the method of steps. The "true values" are $a = -2$, $b = 3$, and $r = 1$. The equation and the results of the iteration are given below:

$$\begin{cases} \dot{x}(t) = ax(t) + bx(t-r), & t \geq 0 \\ x(t) = t + 1, & t \leq 0 \end{cases}$$

iterate	a	b	r	error
0	-4.000	7.000	2.000	3.379
1	-0.815	3.537	1.184	2.968
2	-1.596	3.342	1.122	0.775
3	-2.403	3.713	1.002	0.188
4	-2.250	3.361	1.015	0.094
5	-2.352	3.483	1.006	0.093

The convergence of the states is illustrated in Figure 2.

Example 6.3. This case illustrates the effect of a forcing function on the state equation. The nonhomogeneous delay-differential equation

$$\begin{cases} \dot{x}(t) = ax(t) + bx(t-r) + u(t), & t \geq 0 \\ x(t) = t + 1, & t \leq 0 \end{cases}$$

where

$$u(t) = \begin{cases} 0, & t < 0.1 \\ 1, & t \geq 0.1 \end{cases}$$

is solved in closed form by the method of steps with parameter values $a = -2$, $b = 3$, $r = 1$ as in Example 6.2. The results of the parameter estimation algorithm are given below:

iterate	a	b	r	error
0	-4.000	7.000	2.000	4.0527
1	1.022	3.165	1.140	39.2657
2	-2.637	23.652	1.168	24.9577
3	-5.979	28.631	1.141	11.6964
4	-8.034	23.250	1.118	3.5425
5	-5.167	5.417	1.028	2.0471
6	-1.239	4.195	1.008	4.8981
7	-2.861	6.222	1.005	1.8930
8	-2.485	3.795	0.998	0.0819
9	-2.115	3.201	1.013	0.0724
10	-2.247	3.380	0.998	0.0691

The results are similar to those of Example 6.3, except that the solution has become somewhat more sensitive to the coefficients.

Example 6.4. This example indicates the ability of the algorithm to estimate two unknown delays. The algorithm converges rapidly from a relatively poor initial guess. The "true values" are $r_1 = 1.0$ and $r_2 = 2.0$. The equation and the results of the parameter estimation algorithm are given below and the convergence of the states to ten data points on the interval $[0,3]$ is illustrated in Figure 3.

$$\begin{cases} \dot{x}(t) = -x(t) + x(t-r_1) - x(t-r_2), & t \geq 0 \\ x(t) = t + 1, & t \leq 0 \end{cases}$$

iterate	r_1	r_2	error
0	0.600	4.000	7.500
1	1.569	3.216	2.295
2	1.146	2.100	0.100
3	0.977	1.998	0.034
4	0.978	2.003	0.032

Example 6.5. The equation and data for this example are the same as in Example 6.4. In this case the initial guess reverses the order of the "true" delay values. The results of this iteration are given below and coverage of the states on the interval [0,3] is illustrated in Figure 4.

iterate	r_1	r_2	error
0	2.000	1.000	2.460
1	0.483	1.151	1.379
2	1.561	2.014	0.788
3	1.100	2.072	0.077
4	0.980	2.002	0.033

Example 6.6. In this case the algorithm is asked to estimate parameters in a delay model of a system with no delay. Ten data points on the interval [0,2] are computed from the exponential solution of

$$\begin{cases} \dot{x}(t) = -2x(t) \\ x(0) = 1 \end{cases}$$

and the algorithm is asked to estimate unknown parameters in the system

$$\begin{cases} \dot{x}(t) = ax(t) + bx(t-r), & t \geq 0 \\ x(t) = t + 1, & t \leq 0 \end{cases}$$

The first four iterations are given below:

iterate	a	b	r	error
0	-3.000	3.000	2.000	1.2577
1	-3.060	-0.637	1.947	0.2551
2	-1.687	0.235	1.981	0.1144
3	-1.967	0.025	1.985	0.0110
4	-2.000	0.000	1.986	0.0001

On the fifth iteration the algorithm aborted when it was asked to invert a

nearly singular matrix. This reflects the fact that at the true parameter values the state is completely insensitive to the delay.

Example 6.7. This case is the same as the previous example except that the data is taken from the closed form solution of the nonhomogeneous undelayed equation

$$\begin{cases} \dot{x}(t) = -2x(t) + u(t) \\ x(0) = 1 \end{cases}$$

where u is the same step function as in Example 6.3. The results are similar to those of the previous example.

iterate	a	b	r	error
0	-3.000	3.000	2.000	1.3135
1	-2.848	0.099	1.804	0.5121
2	-1.841	0.138	2.401	0.0811
3	-1.971	0.003	2.508	0.0197

Example 6.8. In this example we consider the second-order equation

$$\begin{cases} \frac{d^2 x}{dt^2}(t) + \omega^2 x(t) + a_0 \frac{dx}{dt}(t-r) + a_1 x(t-r) = u(t), \quad t \geq 0, \\ x(t) = 1, \quad t < 0, \end{cases}$$

where $u(t)$ is the step function of Example 6.3. This equation models a harmonic oscillator with retarded damping and restoring forces. In [13] a quasilinearization algorithm is used to estimate coefficients in this equation. The methods of this paper allow the delay r to be added to the set of unknown parameters. For this example the averaging method was used to compute "data" values for the parameter estimation algorithm with "true" values of $\omega = 6$, $a_0 = 2.5$, $a_1 = 9$, and $r = 1$. The results of the iterative algorithm are given below and the convergence of the states (displacement and velocity) on the interval $[0, 2]$ is illustrated in Figures 5 and 6.

iterate	ω	a_0	a_1	r	error
0	4.100	4.600	6.300	1.500	15.212
1	5.073	6.025	-8.338	0.918	15.181
2	6.705	4.710	-0.682	1.524	12.389
3	6.188	-14.677	-4.838	1.102	31.950
4	5.902	12.347	8.396	1.068	25.234
5	5.964	2.994	8.980	1.061	2.186
6	5.995	2.416	9.016	1.004	0.344
7	6.000	2.503	8.999	1.000	0.007

References

- [1] H. T. Banks, *Identification of nonlinear delay systems using spline methods*, International Conference on Nonlinear Phenomena in the Mathematical Sciences, University of Texas, Arlington, 1980
- [2] H. T. Banks and J. A. Burns, *Hereditary control problems: numerical methods based on averaging approximations*, SIAM J. Contr. Optim. 16, 169-208 (1978)
- [3] H. T. Banks, J. A. Burns, and E. M. Cliff, *A comparison of numerical methods for identification and optimization problems involving control systems with delays*, LCDS Report No. 79-7, Brown University (1979)
- [4] H. T. Banks, J. A. Burns, and E. M. Cliff, *Parameter estimation and identification for systems with delays*, SIAM J. Contr. Optim. 19, 791-828 (1981)
- [5] H. T. Banks, J. A. Burns, and E. M. Cliff, *Spline-based approximation methods for control and identification of hereditary systems*, International Symposium on Systems Optimization and Analysis, A. Bensoussan and J.L. Lions, eds., Lecture Notes in Control and Info. Sci., Vol. 14, Springer, Heidelberg, 314-320, 1979

- [6] H. T. Banks and P. L. Daniel, *Estimation of delays and other parameters in nonlinear functional equations*, SIAM J. Contr. Optim. 21, 893-915 (1983)
- [7] H. T. Banks and G. M. Groome, Jr., *Convergence theorems for parameter estimation by quasilinearization*, J. Math. Anal. Appl. 42, 91-109 (1973)
- [8] H. T. Banks and F. Kappel, *Spline approximation for functional differential equations*, J. Differential Eqns. 34, 496-522 (1979)
- [9] H. T. Banks and K. Kunisch, *Estimation techniques for distributed parameter systems*, Birkhauser, New York, 1989
- [10] H. T. Banks and K. Kunisch, *Parameter estimation techniques for nonlinear distributed parameter systems*, International Conference on Nonlinear Phenomena in the Mathematical Sciences, University of Texas, Arlington, 1980
- [11] D. W. Brewer, *The differentiability with respect to a parameter of the solution of a linear abstract Cauchy problem*, SIAM J. Math. Anal. 13, 607-620 (1982)
- [12] D. W. Brewer, *Quasi-Newton methods for parameter estimation in functional differential equations*, Proceedings of the 27th IEEE Conference on Decision and Control, Austin, Texas, 806-809, 1988
- [13] J. A. Burns and E. M. Cliff, *An abstract quasi-linearization algorithm for estimating parameters in hereditary systems*, IEEE Trans. Automat. Contr. 25, 126-129 (1980)
- [14] J. A. Burns, T. L. Herdman, and H. Stech, *Linear functional differential equations as semi-groups on product spaces*, SIAM J. Math. Anal. 14, 98-116 (1983)

- [15] P. L. Butzer and H. Berens, *Semi-groups of operators and approximation*, Springer-Verlag, New York, 1967
- [16] J. S. Gibson and L. G. Clark, *Sensitivity analysis of a class of evolution equations*, J. Math. Anal. Appl. 58, 22-31 (1977)
- [17] P. W. Hammer, *Parameter identification in parabolic partial differential equations using quasilinearization*, Ph.D. Thesis, ICAM Report 90-07-01, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1990.
- [18] I. G. Rosen, *Discrete approximation methods for parameter identification in delay systems*, SIAM J. Contr. Optim. 22, 95-120 (1984)
- [19] E. Sachs, *Convergence rates of quasi-Newton algorithms for some nonsmooth optimization problems*, SIAM J. Contr. Optim. 23, 401-418 (1985)

Figure 1

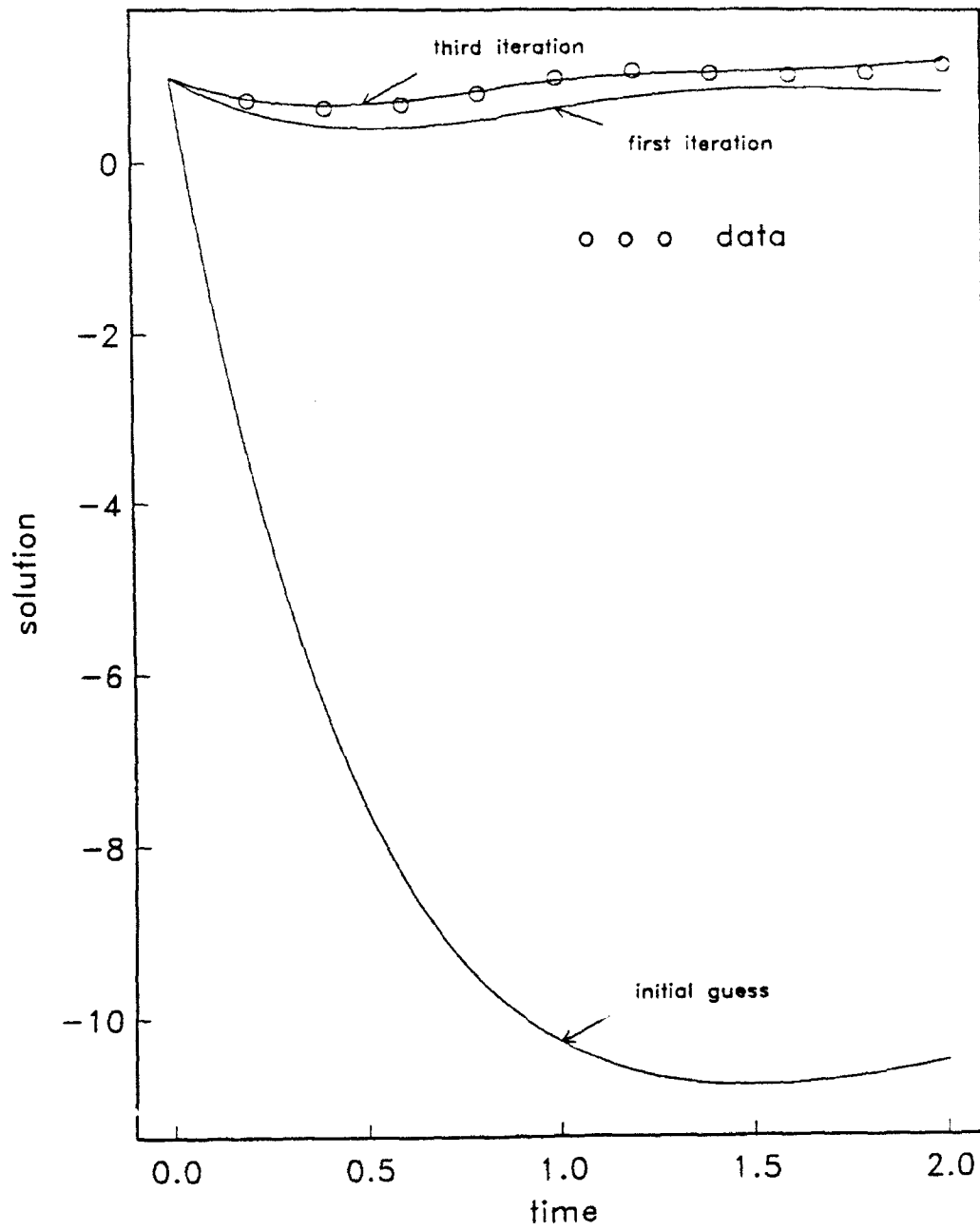


Figure 2

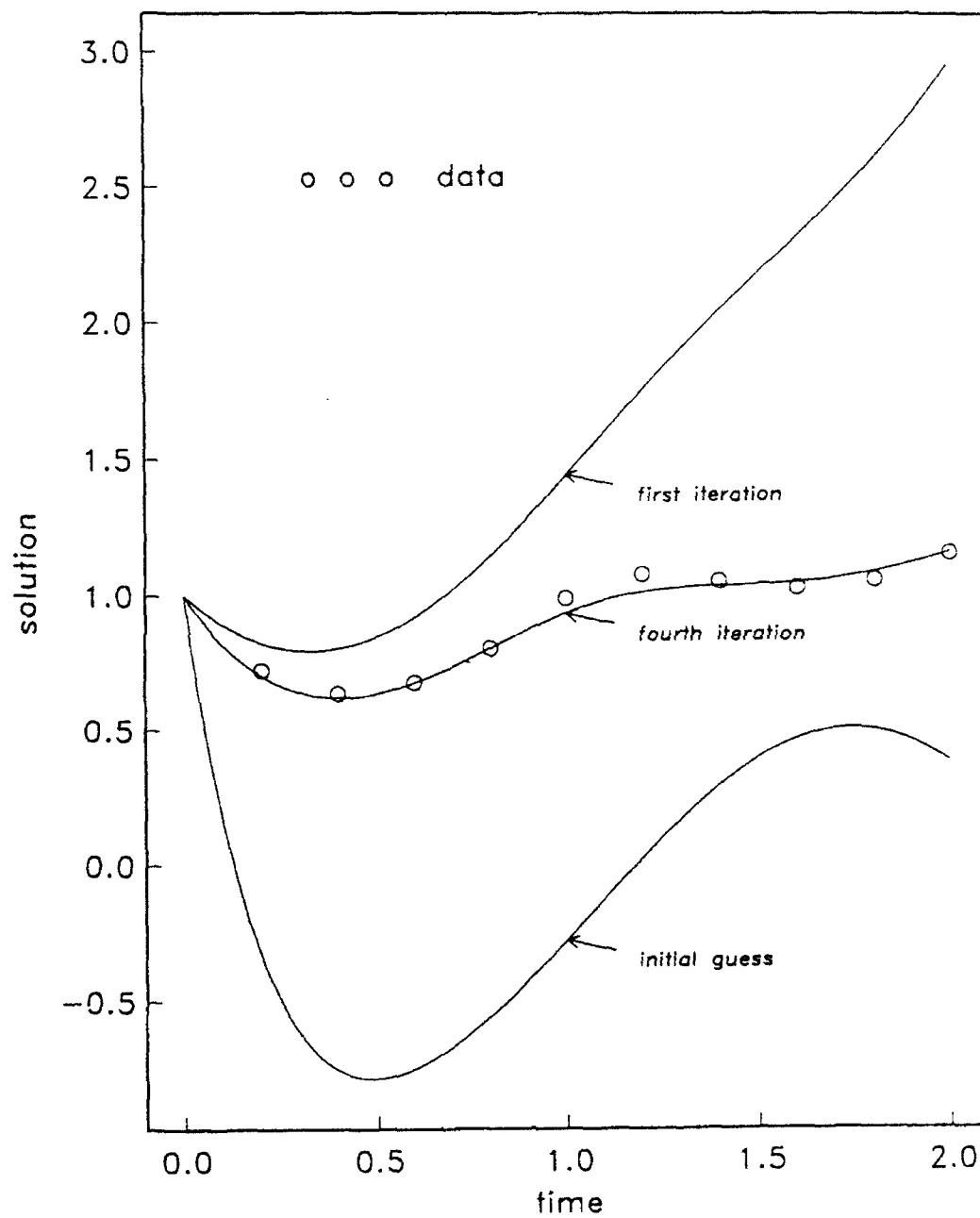


Figure 3

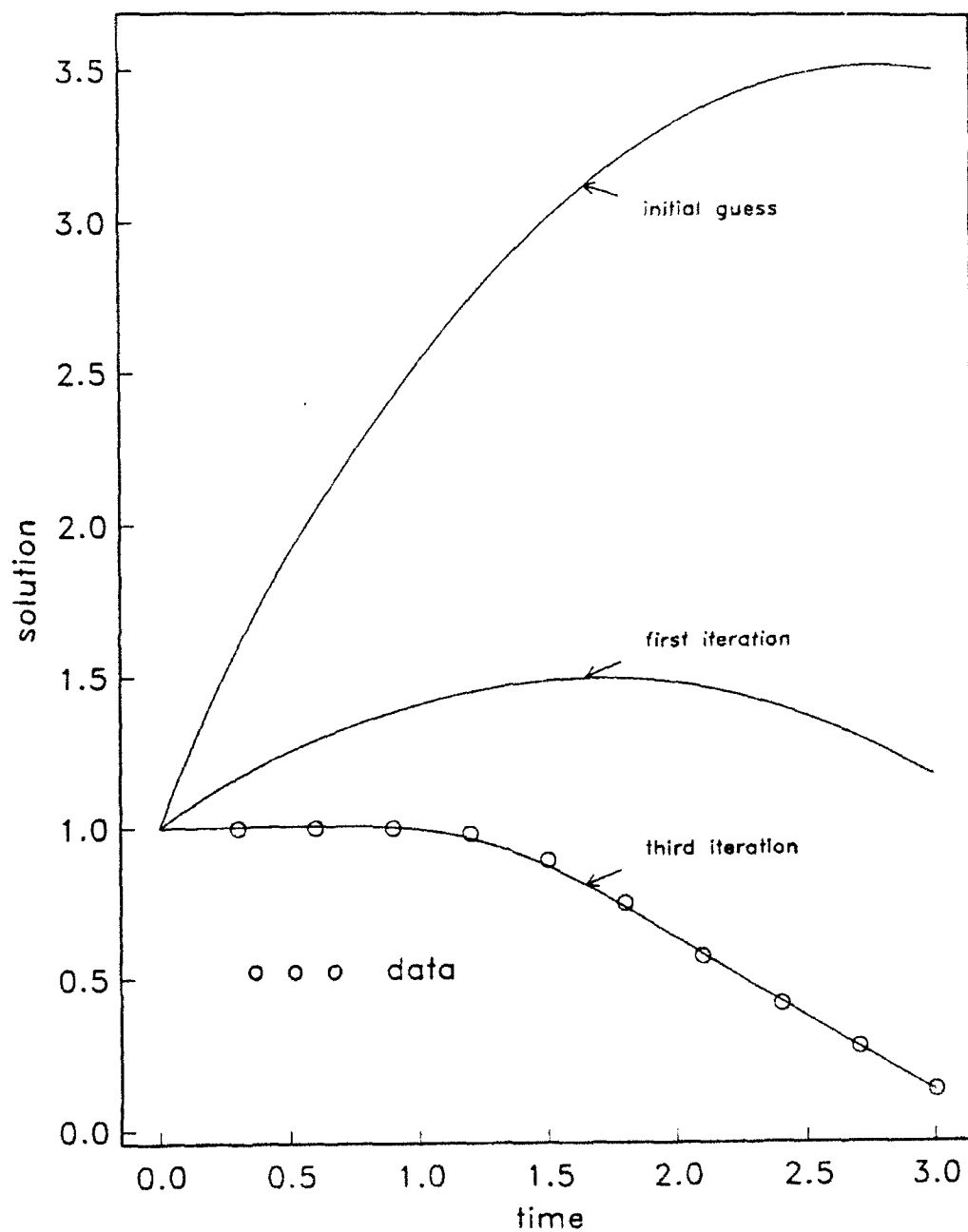


Figure 4

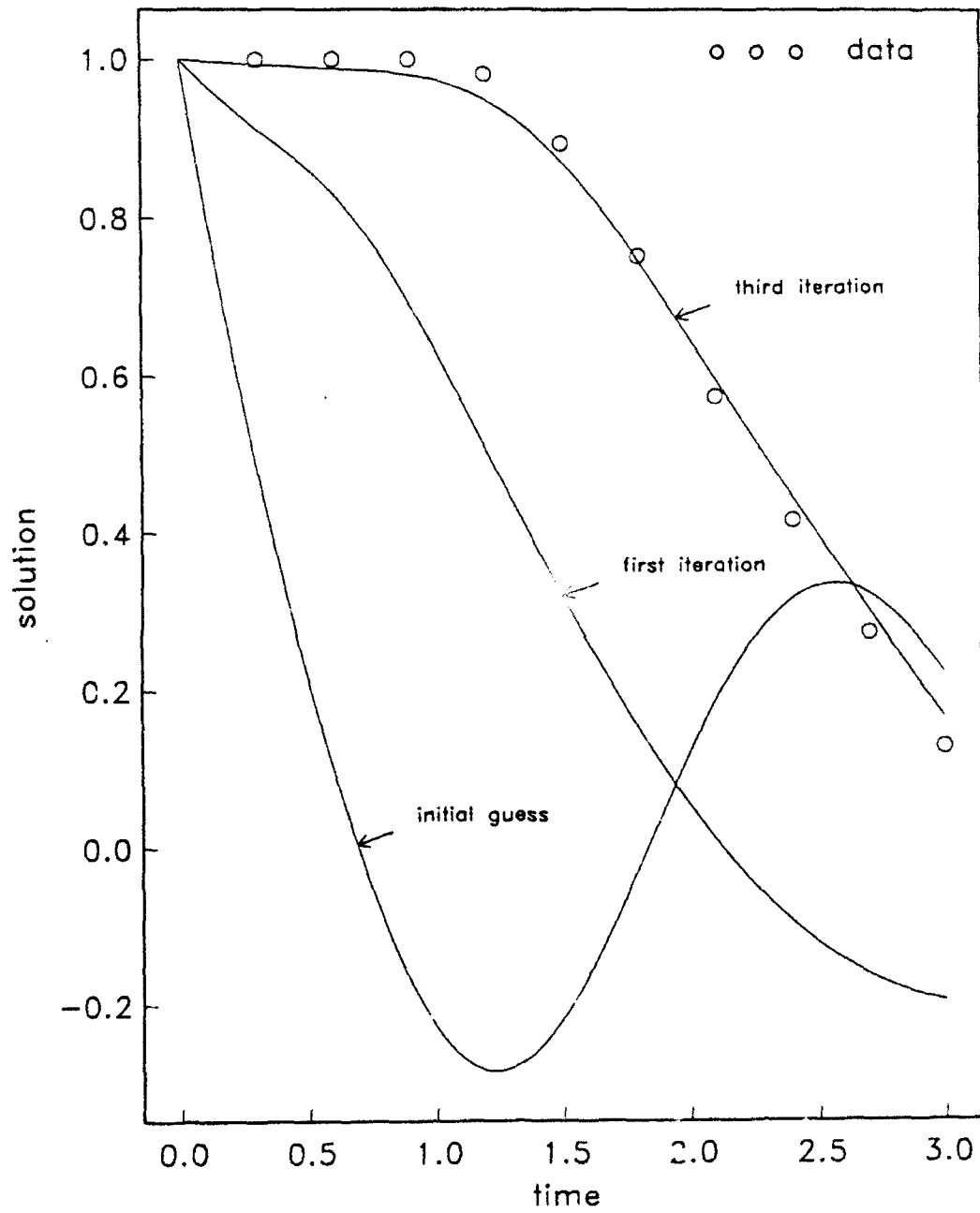


Figure 5

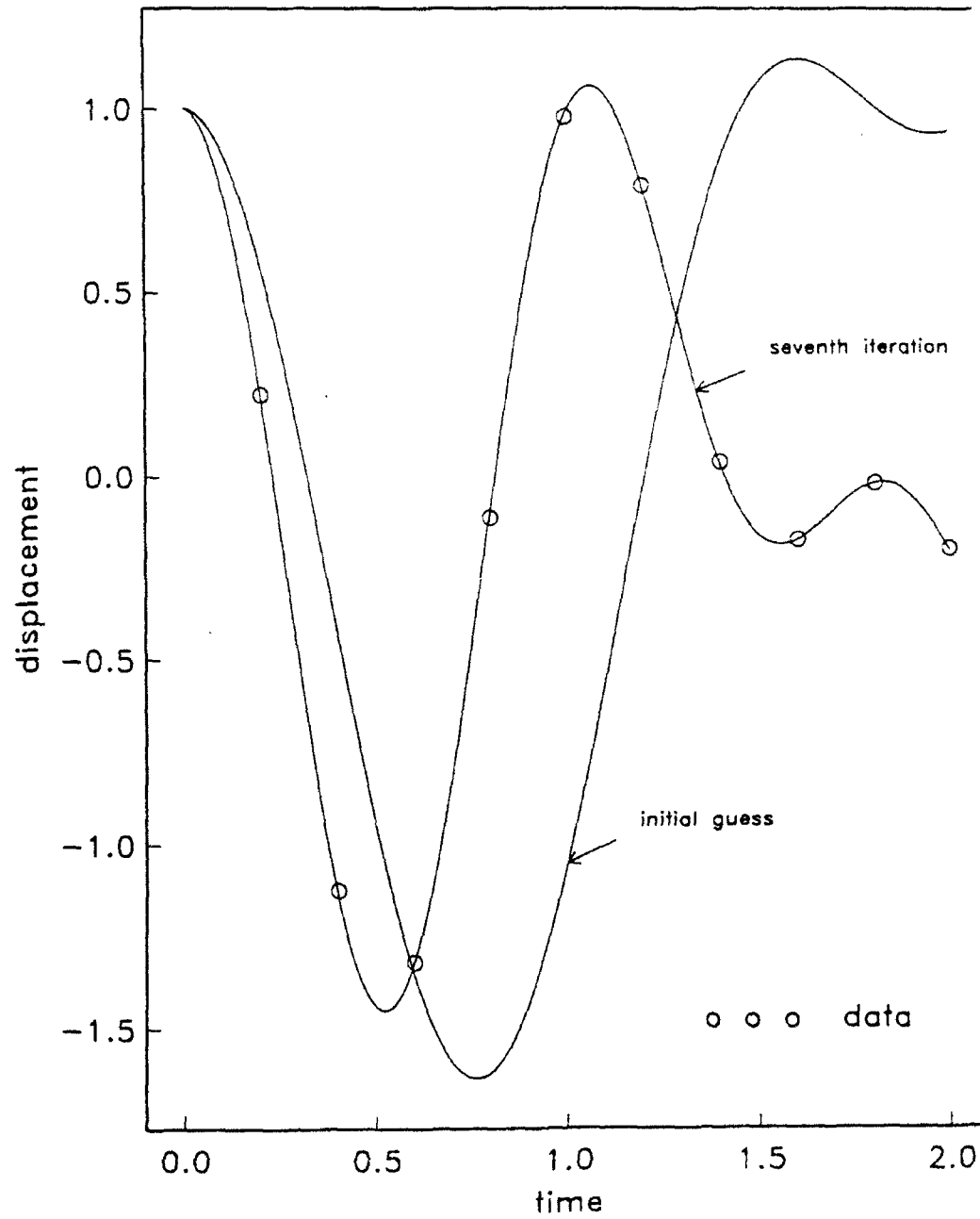
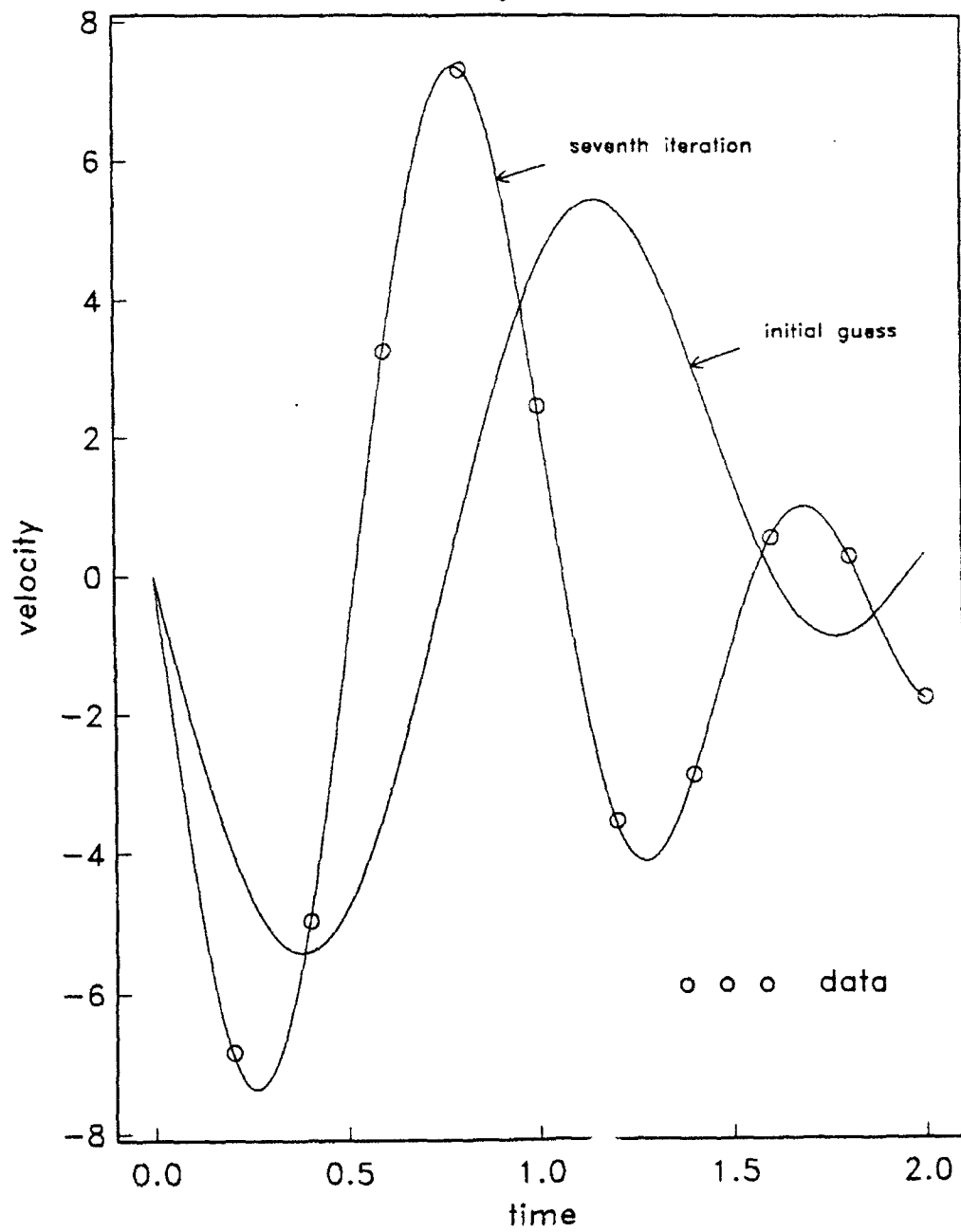


Figure 6



PARAMETER IDENTIFICATION IN A VOLTERRA EQUATION
WITH WEAKLY SINGULAR KERNEL¹

Dennis W. Brewer²

and

Robert K. Powers

Department of Mathematical Sciences
University of Arkansas
Fayetteville, Arkansas 72701

Abstract. We consider identification of parameters in a Volterra integro-differential system with a weakly singular kernel. Such kernels arise in fractional derivative damping models of viscoelastic materials. The Volterra equation is cast in a semigroup setting to establish results on the differentiability of the solution with respect to a parameter. These results are needed for convergence of the identification algorithm. Numerical results are presented.

1. Introduction. In this paper we consider the identification of parameters in a Volterra integro-differential equation with a singular kernel. The equation of interest has the form

$$(1.1) \quad \begin{cases} \dot{w}(t) = Mw(t) + \int_{-\infty}^t K(t-s, p)w(s)ds + F(t), & t \geq 0, \\ w(0) = \eta, \quad w(s) = \phi(s), & s < 0, \end{cases}$$

¹This research was supported by the Air Force Office of Scientific Research under grant AFOSR-89-0472.

²This paper is dedicated with gratitude to my thesis advisor, John A. Nohel, on the occasion of his sixty-fifth birthday.

where M is an $n \times n$ constant matrix, $\eta \in \mathbb{R}^n$, $\phi \in L^1(-\infty, 0; \mathbb{R}^n)$ and $K(\cdot, p)$ is an $n \times n$ singular kernel depending on a parameter p contained in an admissible parameter set. We are particularly interested in a kernel function of the form

$$g(s, p) = \frac{\gamma e^{-\beta s}}{\Gamma(1-\alpha) s^\alpha}, \quad s > 0,$$

where $\Gamma(\cdot)$ denotes the gamma function, $p = (\alpha, \beta, \gamma) \in \mathbb{R}^3$ with $0 \leq \alpha < 1$ and $\beta > 0$. Such kernels arise in the study of fractional derivative models of viscoelastic structures. For a more complete discussion of the origins of this kernel and the viscoelastic models we refer the reader to [12], [18], [13], [15], and in particular to [17] and the extensive bibliography therein.

Banks, et.al. [2] have identified parameters corresponding to β and γ in a similar (but different) model, but assumed that α was known. Torvik and Bagley [1], [18] have estimated the parameter α , but in the Laplace transform domain. In this paper we restrict ourselves to identifying α only, though the theory may be modified to include β and γ as well.

In order to relate equation (1.1) to a (idealized) physical model, consider the longitudinal motions of a uniform bar fixed at both ends with Boltzmann type damping. The governing equation is ([9], [14])

$$(1.2) \quad \begin{cases} \rho u_{tt}(x, t) = \frac{\partial}{\partial x} \left\{ E u_x(x, t) + \frac{\partial}{\partial t} \int_0^t g(t-s) u_x(x, s) ds \right\} \\ + f(x, t), \quad 0 < x < 1, \quad t > 0, \end{cases}$$

with boundary conditions $u(0, t) = 0, \quad u(1, t) = 0,$
and initial conditions $u(x, 0) = d(x), \quad u_t(x, 0) = v(x).$

Here, $u(x, t)$ represents the axial displacement of position x at time t , ρ is the density of the material, E a stiffness parameter, $f(x, t)$ a forcing function, and

$$g(s) = \frac{\gamma e^{-\beta s}}{\Gamma(1-\alpha) s^\alpha}$$

represents a fractional derivative damping term modified to have exponential decay [12].

A common approach to the parameter identification problem [4] is to apply a Galerkin-type approximation scheme to the

beam equation and then incorporate some type of identification algorithm to the approximating system of integro-differential equations. If one applies a Galerkin scheme to equation (1.2) (e.g. using linear splines), one obtains a system of equations of the form

$$(1.3) \quad \left\{ \hat{A} \frac{d^2 v}{dt^2}(t) = \hat{B} v(t) + \hat{C} \frac{d}{dt} \int_0^t g(t-s) v(s) ds + \hat{F}(t) \right. .$$

In this equation \hat{A} , \hat{B} , and \hat{C} represent constant matrices and $v(t)$ and $F(t)$ are vectors of appropriate dimension.

In order to retain the salient features but simplify the analysis in the following sections, we shall consider the following scalar version of (1.3) :

$$(1.4) \quad \left\{ \begin{aligned} \frac{d^2 x}{dt^2}(t) &= ax(t) + \frac{d}{dt} \int_0^t g(t-s)x(s)ds + \hat{f}(t), \\ x(0) &= x_0, \quad \dot{x}(0) = x_1 . \end{aligned} \right.$$

Integrating (1.4) we obtain

$$(1.5) \quad \left\{ \begin{aligned} \dot{x}(t) &= a \int_0^t x(s)ds + \int_0^t g(t-s)x(s)ds + f(t), \\ x(0) &= x_0, \end{aligned} \right.$$

where
$$f(t) = x_1 + \int_0^t \hat{f}(s)ds .$$

Define $z(t) = \int_0^t x(s)ds$, then $\dot{z}(t) = x(t)$ and we obtain the system of integro-differential equations

$$(1.6) \quad \left\{ \begin{aligned} \dot{x}(t) &= az(t) + \int_0^t g(t-s)x(s)ds + f(t), \\ \dot{z}(t) &= x(t), \end{aligned} \right.$$

with $x(0) = x_0, \quad z(0) = 0.$

A standard assumption in viscoelasticity [13] is that the material is in an unstrained state for time $t < 0$. This would correspond to $u(x,s) = 0$ for $s < 0$ in equation (1.2). It follows then that $x(s) = 0$ and $z(s) = 0$ for $s < 0$ in (1.6). If we define $w(t) = \text{col}(x(t), z(t))$, then $w(t)$ satisfies

$$(1.7) \quad \begin{cases} \dot{w}(t) = Mw(t) + \int_0^t K(t-s)w(s)ds + F(t), \\ w(0) = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}, \end{cases}$$

where $M = \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}$, $K(s) = \begin{pmatrix} g(s) & 0 \\ 0 & 0 \end{pmatrix}$, and $F(t) = \begin{pmatrix} f(t) \\ 0 \end{pmatrix}$.

Since $w(s) = \text{col}(0,0)$ for $s < 0$, we may rewrite (1.7) as

$$\begin{aligned} \dot{w}(t) &= Mw(t) + \int_{-\infty}^t K(t-s)w(s)ds + F(t), \\ w(0) &= \begin{pmatrix} x_0 \\ 0 \end{pmatrix}, \quad w(s) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad s < 0, \end{aligned}$$

which is in the form of equation (1.1).

The remainder of the paper is outlined as follows. In Section 2 we review previous results that place equation (1.1) in a semigroup setting in order to establish existence of solutions. Differentiability results needed for the parameter estimation algorithm are then proved. In Section 3 the quasilinearization algorithm used for the identification procedure is discussed along with convergence results. Numerical examples are given in Section 4.

2. The abstract setting. In this section we develop an abstract framework for the Volterra integral equation discussed in the previous section. Namely, we will consider equation (1.1) in the form

$$(2.1) \quad \begin{cases} \dot{w}(t) = Mw(t) + \int_{-\infty}^t K(-s, \alpha)w(t+s)ds + F(t), \quad t > 0, \\ w(0) = \eta, \quad w(s) = \varphi(s), \quad s < 0, \end{cases}$$

where $\eta = \begin{pmatrix} x_0 \\ 0 \end{pmatrix} \in \mathbb{R}^2$, $M = \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}$, $F(t) = \begin{pmatrix} f(t) \\ 0 \end{pmatrix}$, and

$$(2.2) \quad K(s, \alpha) = \frac{\gamma e^{-\beta s}}{\Gamma(1-\alpha)s^\alpha} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad s > 0.$$

We assume β is a positive constant and $0 \leq \alpha < 1$. By a solution of (2.1) we mean a function $w: (-\infty, \infty) \rightarrow \mathbb{R}^2$ such that w is absolutely continuous (A.C.) on $[0, \infty)$ and satisfies the integral equation a.e. on $[0, \infty)$, $w(0) = \eta$, and $w(s) = \varphi(s)$ a.e. on $(-\infty, 0]$.

Our semigroup formulation follows the construction in [5] as further developed in [10] and [11]. Define the product space $X = \mathbb{R}^2 \times L^1(-\infty, 0)$ with norm $\|(\eta, \varphi)\|_X = |\eta| + \|\varphi\|_{L^1(-\infty, 0)}$. Consider the homogeneous equation

$$(2.3) \quad \begin{cases} \dot{y}(t) = My(t) + \int_{-\infty}^t K(-s, \alpha) y(t+s) ds, & t > 0, \\ y(0) = \eta, \quad y(s) = \varphi(s), & s < 0. \end{cases}$$

Then for each pair $(\eta, \varphi) \in X$, (2.3) has a unique solution, and moreover the mapping $S(t, \alpha)(\eta, \varphi) = (y(t), y_t(\cdot))$ defines a C_0 -semigroup on X . Here we have used the notation $y_t(s) = y(t+s)$, $t \geq 0$, $s < 0$.

Fix $\varepsilon \in (0, 1)$ and define the parameter set $P = [0, 1-\varepsilon]$. Then it is readily seen from (2.2) that there is a constant C , independent of α , such that

$$(2.4) \quad \int_0^\infty |K(s, \alpha)| ds \leq C \quad \text{for all } \alpha \in P.$$

Under this condition it is shown in [10] and [11] that the semigroup $S(t, \alpha)$ is generated by a closed and densely-defined operator $\mathcal{A}(\alpha)$ defined by

$$\text{Dom}(\mathcal{A}(\alpha)) \equiv D = \{(\eta, \varphi) \in X: \varphi \text{ is A.C. on compact subsets of } (-\infty, 0], \dot{\varphi} \in L^1(-\infty, 0), \varphi(0) = \eta\}$$

and

$$\mathcal{A}(\alpha)(\eta, \varphi) = (M\eta + \int_{-\infty}^0 K(-s, \alpha) \varphi(s) ds, \dot{\varphi}) \text{ for } (\eta, \varphi) \in D.$$

Our task is to show that the solution $w(t, \alpha) \equiv w(t)$ of (2.1) is differentiable with respect to α and that this derivative is sufficiently smooth to establish the local convergence of the algorithm defined in Section 3. This involves verifying the conditions in the semigroup setting established in [6] and [8]. We therefore assume in what follows that the reader has these papers in hand.

Since we are interested in dependence on α , we write

$A(\alpha) = A + B(\alpha)$ where A is independent of α and

$$(2.5) \quad B(\alpha)(\eta, \varphi) = \left(\int_{-\infty}^0 K(-s, \alpha) \varphi(s) ds, 0 \right), \quad (\eta, \varphi) \in D.$$

Note that the range of $B(\alpha)$ is the finite-dimensional space $Y = \mathbb{R}^2 \times \{0\}$.

Fix $y_0 \in X$, $\alpha_0 \in P$, and $T > 0$. Then the differentiability with respect to α at α_0 of the solution $y(t, \alpha)$ of (2.3) is a consequence of the following theorem.

Theorem 2.1. For every $t \in [0, T]$, $S(t, \alpha)y_0$ as defined above is Frechét differentiable with respect to α at α_0 and its derivative is given by

$$D_{\alpha} S(t, \alpha_0)y_0 = \int_0^t S(t-s, \alpha_0) [D_{\alpha} \mathcal{F}(\alpha_0)y_0](s) ds, \quad 0 \leq t \leq T,$$

where $\mathcal{F}(\alpha)$ is for each $\alpha \in P$ a mapping from X into $L^1(0, T; X)$ defined by

$$(2.6) \quad [\mathcal{F}(\alpha)y_0](t) = \left(\int_{-\infty}^0 K(-s, \alpha) y(t+s) ds, 0 \right)$$

for $y_0 \in X$, $0 \leq t \leq T$. Recall that y is the solution of (2.3) with $(\eta, \varphi) = y_0$.

Proof: This result is proved in [7] for a general Volterra kernel $K(s, \alpha)$ satisfying condition (2.4) under the following hypothesis:

(2.7) the mapping $\alpha \rightarrow K(\cdot, \alpha)$ from P into $L^1(0, \infty)$ is Frechét differentiable with respect to α at α_0 .

Recall that $K(s, \alpha) = g(s, \alpha) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ where

$$g(s, \alpha) = \frac{\gamma e^{-\beta s}}{\Gamma(1-\alpha) s^{\alpha}}, \quad s > 0, \quad \alpha \in P.$$

Let ' denote differentiation with respect to α . Then computation shows that g' and g'' are of the form

$$g'(s, \alpha) = g_1(\alpha) (\ln s) e^{-\beta s} s^{-\alpha} + g_2(\alpha) e^{-\beta s} s^{-\alpha}$$

and

$$g''(s, \alpha) = g_3(\alpha) (\ln s)^2 e^{-\beta s} s^{-\alpha} + g_4(\alpha) (\ln s) e^{-\beta s} s^{-\alpha} + g_5(\alpha) e^{-\beta s} s^{-\alpha}$$

where g_1, \dots, g_5 are continuous functions of α on P which can be explicitly calculated in terms of the gamma function and its derivatives. Important properties of g' and g'' for our purposes are that there are functions $\psi_1, \psi_2 \in L^1(0, \infty)$ such that

$$(2.8) \quad |g'(s, \alpha)| \leq \psi_1(s) \quad \text{for } s > 0, \alpha \in P,$$

and

$$(2.9) \quad |g''(s, \alpha)| \leq \psi_2(s) \quad \text{for } s > 0, \alpha \in P.$$

For example, one can take

$$\psi_1(s) = \begin{cases} (M_1 |\ln s| e^{-\beta s} + M_2 e^{-\beta s}) / s^{1-\varepsilon}, & \text{for } 0 < s < 1 \\ M_1 |\ln s| e^{-\beta s} + M_2 e^{-\beta s}, & \text{for } s \geq 1, \end{cases}$$

where M_1 and M_2 are upper bounds on P of $|g_1(\alpha)|$ and $|g_2(\alpha)|$, respectively. There is an analogous expression for $\psi_2(s)$. Therefore, by Taylor's theorem with remainder, we obtain

$$\begin{aligned} & |K(s, \alpha + h) - K(s, \alpha) - K'(s, \alpha)h| \\ &= |g(s, \alpha + h) - g(s, \alpha) - g'(s, \alpha)h| \\ &= |g''(s, \xi_1(s))h^2/2| \\ &\leq \psi_2(s) |h|^2/2 \end{aligned}$$

for $s > 0$, $\alpha, \alpha + h \in P$, and $\xi_1(s)$ between α and $\alpha + h$.

Integrating this inequality over $(0, \infty)$ and using $\psi_2 \in L^1(0, \infty)$ yields (2.7) and completes the proof of Theorem 2.1.

A sufficient smoothness property for the local convergence of the parameter estimation algorithm is established in the following theorem.

Theorem 2.2. For every $t \in [0, T]$, $\alpha^* \in P$ and $y_0 \in X$, $D_\alpha S(t, \alpha)y_0$ is strongly locally Lipschitz continuous with respect to α at α^* .

Proof: The proof relies on Lemma 3.3 of [8]. We must show that hypotheses (H11) and (H12) of that paper hold in this application. By definition (2.6), (H11) requires that there exist constants $K_1, \delta_1 > 0$ such that

$$(2.10) \quad \int_0^T \left| \int_{-\infty}^0 (K(-s, \alpha+h) - K(-s, \alpha)) y(t+s) ds \right| dt \leq K_1 |h| |\eta|$$

for $|h| \leq \delta_1$, where y is the solution of (2.3) with $\alpha = \alpha^*$ and $\varphi \equiv 0$. Note that since $(y(t), y_t) = S(t, \alpha_0)(\eta, 0)$ we have $|y(t)| \leq M_1 e^{\omega t} |\eta|$ for $t \geq 0$, and $y(t) = \varphi(t) = 0$ for $t < 0$. It is shown in [7] that the constants M_1 and ω may be taken independently of $\alpha \in P$. Therefore, by Fubini's theorem and the mean value theorem we obtain

$$\begin{aligned} & \int_0^T \int_{-\infty}^0 |(K(-s, \alpha+h) - K(-s, \alpha)) y(t+s)| ds dt \\ &= \int_{-\infty}^0 |K(-s, \alpha+h) - K(-s, \alpha)| \int_s^{T+s} |y(t)| dt ds \\ &\leq \int_{-\infty}^0 |K(-s, \alpha+h) - K(-s, \alpha)| \int_0^T |y(t)| dt ds \\ &\leq M_1 T e^{\omega T} |\eta| \int_{-\infty}^0 |K(-s, \alpha+h) - K(-s, \alpha)| ds \\ &\leq M_1 T e^{\omega T} |\eta| |h| \int_{-\infty}^0 \psi_1(-s) ds \end{aligned}$$

for $\alpha, \alpha + h \in P$. Here we have used (2.8) in the last inequality. Since $\psi_1 \in L^1(0, \infty)$, this establishes (2.10).

Hypothesis (H12) of [8] requires the Lipschitz continuity of the derivative with respect to α of the mapping $\mathcal{F}(\alpha)$ defined by (2.6). For brevity, we denote the value of this derivative at $\alpha \in P$ by $D\mathcal{F}(\alpha)$. The existence of $D\mathcal{F}(\alpha)$ was shown in Theorem 2.1 and from the proof of that theorem we have the formula

$$[D\mathcal{F}(\alpha)h](t) = \left(\int_{-\infty}^0 [K'(-s, \alpha)h] y(t+s) ds, 0 \right)$$

for $0 \leq t \leq T$, $h \in \mathbb{R}$, $\alpha \in P$, where y is the solution of (2.3)

with $\alpha = \alpha_0$ and $(\eta, \varphi) = y_0$. Recall that ' denotes differentiation with respect to α . The local Lipschitz continuity of $D\mathcal{F}(\alpha)$ at a point $\alpha = \alpha^* \in P$ now follows from estimates similar to those used to establish (H11) but with ψ_2 in place of ψ_1 . This completes the proof of Theorem 2.2.

We now turn our attention to solutions of the nonhomogeneous equation (2.1). It is well-known that a mild solution to this equation is given by the variation of constants formula

$$(w(t), w_t) = S(t, \alpha)(\eta, \varphi) + Q(t, \alpha)$$

where

$$(2.11) \quad Q(t, \alpha) = \int_0^t S(t-s, \alpha)(F(s), 0) ds.$$

It remains, therefore, to consider the existence and smoothness of the derivative of $Q(t, \alpha)$ with respect to α . We again appeal to [8] where these properties of $Q(t, \alpha)$ are demonstrated by considering similar properties of the mapping $\mathcal{S}(\alpha): L^1(0, T; X) \rightarrow L^1(0, T; X)$ defined by

$$[\mathcal{S}(\alpha)v](t) = \int_0^t \mathcal{B}(\alpha)S(t-s, \alpha_0)v(s) ds$$

for $v \in L^1(0, T; X)$, $\alpha \in P$, α_0 fixed. Note that if $v(t) = (F(t), 0)$ for some $F \in L^1(0, T)$, then

$$(w(t), w_t) = \int_0^t S(t-s, \alpha_0)v(s) ds$$

where w is a mild solution of (2.1) with $(\eta, \varphi) = (0, 0)$. Since $\mathcal{B}(\alpha)$ is a difference of closed operators,

$$[\mathcal{S}(\alpha)v](t) = \mathcal{B}(\alpha)(w(t), w_t).$$

Therefore, using definition (2.5) we obtain in this setting that

$$(2.12) \quad [\mathcal{S}(\alpha)v](t) = \left(\int_{-\infty}^0 K(-s, \alpha)w(t+s) ds, 0 \right)$$

where $v(t) = (F(t), 0)$ and w is the solution of (2.1) with $\alpha = \alpha_0$ and $(\eta, \varphi) = (0, 0)$. Here we assume F is sufficiently

smooth for the solution w to exist in the strong sense defined earlier.

Comparing definitions (2.6) and (2.12), we see that properties of $\mathcal{G}(\alpha)$ with respect to α can be proven in the same way as the corresponding properties of $\mathcal{F}(\alpha)$ using the solution of (2.1) in place of the solution of (2.3) at a fixed $\alpha_0 \in P$.

For this reason the following theorems, which are consequences of Lemmas 3.2 and 3.4 of [8], are stated without proof.

Theorem 2.3. For $F \in L^1(0, \infty)$, let $Q(t, \alpha)$ be defined by (2.11). Then for every $t \in [0, T]$ and $\alpha_0 \in P$, $Q(t, \alpha)$ is Frechét differentiable with respect to α at α_0 and this derivative is given by the formula

$$D_\alpha Q(t, \alpha_0) = \int_0^t S(t-s, \alpha_0) [D_\alpha \mathcal{G}(\alpha_0) v](s) ds$$

where $v(t) = (F(t), 0)$ and $\mathcal{G}(\alpha)$ is defined by (2.12).

Theorem 2.4. Suppose the hypotheses of Theorem 2.3 hold. Then the mapping $D_\alpha Q(t, \alpha)$ is locally Lipschitz continuous with respect to α at every $\alpha^* \in P$.

3. The algorithm. In this section we define a parameter estimation algorithm based on quasilinearization and state some local convergence results. For later adaptation we develop the algorithm for the case $\alpha \in P \subset \mathbb{R}^n$ with canonical basis e_i , $i = 1, 2, \dots, n$. In Section 4 the algorithm is applied in the case $n = 1$. The definitions and theorems stated here may also be found in [8], but are included here for completeness.

Using the notation of the previous section, let $y_0 = (\eta, \varphi) \in X$ and $\alpha \in P$. Let C be a bounded linear mapping from X into a finite-dimensional space \mathbb{R}^ℓ , and define

$$w(t, \alpha) = C[S(t, \alpha)y_0 + Q(t, \alpha)].$$

The parameter estimation algorithm is related to the following optimization problem.

Problem 3.1. Let $w_j \in \mathbb{R}^\ell$, $j = 1, 2, \dots, m$ be data values taken at times $t_j \in [0, T]$, $j = 1, 2, \dots, m$, respectively. For $\alpha \in P$ define the quadratic cost function

$$J(\alpha) = \sum_{j=1}^m |w(t_j, \alpha) - w_j|^2.$$

Find $\alpha^* \in P$ such that $J(\alpha^*) \leq J(\alpha)$ for all $\alpha \in P$.

The quasilinearization algorithm method defines a recursive algorithm whose fixed point is a local solution of Problem 3.1. A more complete exposition is given in [3]. Given an initial guess $\alpha_0 \in P$ define

$$\alpha_{k+1} = \ell(\alpha_k), \quad k = 0, 1, 2, \dots$$

where

$$\ell(\alpha) = \alpha - [D(\alpha)]^{-1}b(\alpha)$$

$$D(\alpha) = \sum_{j=1}^m M^T(t_j, \alpha) M(t_j, \alpha)$$

$$b(\alpha) = \sum_{j=1}^m M^T(t_j, \alpha) [w(t_j, \alpha) - w_j]$$

and the matrix $M(t, \alpha)$ has its i^{th} column $M^i(t, \alpha)$ given by

$$M^i(t, \alpha) = CD_\alpha[S(t, \alpha)y_0 + Q(t, \alpha)]e_i, \quad i = 1, 2, \dots, n.$$

The following theorems are typical of quasilinearization methods. Their proofs may be found in [8]. We obtain superlinear convergence when there is an exact fit to data (Theorem 3.1) and linear convergence in the presence of error (Theorem 3.2).

Theorem 3.1. Suppose the conditions of the previous section are satisfied. Moreover, assume $[D(\alpha)]^{-1}$ exists, $\ell(\alpha^*) = \alpha^*$, and $J(\alpha^*) = 0$. Then for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$|\ell(\alpha) - \ell(\alpha^*)| \leq \epsilon |\alpha - \alpha^*|$$

for $|\alpha - \alpha^*| \leq \delta$. In particular, there is a neighborhood \mathcal{U} of α^* such that $\alpha_k \rightarrow \alpha^*$ as $k \rightarrow \infty$ whenever $\alpha_0 \in \mathcal{U}$.

The following theorem does not require an exact fit to data but does place some technical restrictions on the behaviour of the matrix $M(t, \alpha)$ near α^* . Note that under the conditions of Theorem 2.2 there exists number $\bar{\delta} > 0$ such that for $0 < \delta < \bar{\delta}$ there exists a constant $K(\delta)$ such that

$$\sum_{j=1}^m |M^T(t_j, \alpha) - M^T(t_j, \alpha^*)| \leq K(\delta) |\alpha - \alpha^*|$$

for $|\alpha - \alpha^*| \leq \delta$. Let $K^* = \limsup_{\delta \downarrow 0} K(\delta)$ and define

$$\lambda^* = K^* |D(\alpha^*)|^{-1} \max_j |w(t_j, \alpha^*) - w_j|.$$

Theorem 3.2. Suppose the conditions of the previous section are satisfied. Moreover, assume $[D(\alpha^*)]^{-1}$ exists and $f(\alpha^*) = \alpha^*$. Let λ^* be defined as above and assume $\lambda^* < 1$. Then there exists $\delta^* > 0$ such that

$$|f(\alpha) - f(\alpha^*)| \leq \lambda^* |\alpha - \alpha^*|$$

for $|\alpha - \alpha^*| \leq \delta^*$. In particular, $\alpha_k \rightarrow \alpha^*$ as $k \rightarrow \infty$ whenever $|\alpha_0 - \alpha^*| \leq \delta^*$.

4. Numerical results. In this section we present several examples that illustrate the ideas discussed in the previous sections. Recall the identification problem: given observations w_j at times t_j , $j = 1, 2, \dots, m$, determine $\alpha \in [0, 1]$ that minimizes the cost functional

$$J(\alpha) = \sum_{j=1}^m (x(t_j) - w_j)^2$$

where $x(t)$ satisfies

$$(4.1) \quad \begin{cases} \dot{x}(t) = a \int_0^t x(s) ds + \frac{\gamma}{\Gamma(1-\alpha)} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} x(s) ds + f(t), \\ x(0) = x_0. \end{cases}$$

The quasilinearization algorithm requires that we solve (4.1) along with its sensitivity equation which has the form

$$(4.2) \quad \left\{ \begin{aligned} \dot{x}_\alpha(t) &= a \int_0^t x_\alpha(s) ds + \frac{\gamma}{\Gamma(1-\alpha)} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} x_\alpha(s) ds \\ &+ \frac{\gamma \Gamma'(1-\alpha)}{\Gamma(1-\alpha)^2} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} x(s) ds \\ &- \frac{\gamma}{\Gamma(1-\alpha)} \int_0^t \frac{\ln(t-s) e^{-\beta(t-s)}}{(t-s)^\alpha} x(s) ds, \\ x_\alpha(0) &= 0, \end{aligned} \right.$$

where $x_\alpha(t) = \frac{\partial x}{\partial \alpha}(t)$. The zero initial condition reflects the fact that the value $x(0) = x_0$ is independent of the parameter α .

The implementation of the identification scheme begins with an initial guess for α . Equations (4.1) and (4.2) are integrated using this initial value, then $x(t)$ and $x_\alpha(t)$ are used to give an updated estimate of the parameter. For this particular problem the quasilinearization algorithm updates the current estimate α_k according to

$$\alpha_{k+1} = \alpha_k - \frac{\sum_{j=1}^m (x(t_j) - w_j) x_\alpha(t_j)}{\sum_{j=1}^m (x_\alpha(t_j))^2}$$

where $x(t)$ and $x_\alpha(t)$ are the solutions of (4.1)-(4.2) computed for $\alpha = \alpha_k$. In order to numerically integrate the state and sensitivity equations we first convert (4.1)-(4.2) to integral equations via the substitution $z(t) = x(t)$ and $z_\alpha(t) = x_\alpha(t)$.

Then one has the 4×4 system of integral equations consisting of (4.1)-(4.2) with the above substitutions coupled with

$$(4.3) \quad \left\{ \begin{aligned} x(t) &= x(t_p) + \int_{t_p}^t z(s) ds, \\ x_\alpha(t) &= x_\alpha(t_p) + \int_{t_p}^t z_\alpha(s) ds, \end{aligned} \right.$$

where $t_p \in [0, t)$ is determined by the approximation scheme.

The solution of the system of the integral equations is then approximated by applying a product integration method based on Simpson's rule to the singular integral terms, and Simpson's rule to all other integrals. For a description of product integration methods we refer the interested reader to [16].

In each of the following examples we numerically solve (4.1)-(4.3) in time on the interval $[0, 1]$. Define $t_j = j/N$, $j = 0, 1, \dots, N$. The numerical integration scheme then computes values for $x(t_j)$. Examples (4.1) and (4.2) presented here are computed using a value of $N = 50$, and Example (4.3) is computed using $N = 200$. In each case 5 data points located at $t = .2, .4, .6, .8$, and 1.0 are used in the identification procedure. The true values of $x(t)$ in all of the Figures (4.1)-(4.5) are denoted by \times .

Example 4.1. In this example we set the values of a , β , and γ to $1.$, $1.$, and $5.$, respectively. The parameter value to be identified is $\alpha = \frac{1}{2}$, and the nonhomogeneous term $f(t)$ is

$$f(t) = e^{-t} (1024t^3 + 2176t + 1792) - 1825 \\ - \frac{e^{-t} t^{.5}}{\Gamma(.5)} \left(\frac{32768}{315} t^4 - \frac{8192}{35} t^3 + \frac{512}{3} t^2 - \frac{128}{3} t + 2 \right)$$

The true solution is $x(t) = e^{-t} T_4(2t - 1)$ where $T_4(s)$ is the Chebyshev polynomial of degree 4 on $-1 \leq s \leq 1$. Tables (4.1) and (4.2) contain the results for two computer runs, one for an initial α of $\alpha_0 = .9$, and the other for $\alpha_0 = .2$. The sequence of α_k values, their corresponding costs $J(\alpha_k)$, and the values of the state $x(t)$ at time $t = 1$ are included to illustrate the convergence. The true value of $x(1)$ is $.3678794$. Note that in each case once an estimate of α is greater than $.5$, then the sequence of iterates converges monotonically down to the true value. This is a characteristic of all of our computer simulations, and seems to indicate that it is better to choose an initial value of α that is high instead of low. In fact, for all simulations another characteristic is that if the initial choice is excessively low, then the next estimate of α is greater than 1 , and the integral becomes undefined. For this particular example, $\alpha_0 = .1$ resulted in a value greater than 1 for the next iterate and the integration scheme broke down. However, though not shown here, some examples ran successfully even with a negative initial value for α . Figures (4.1)-(4.2) show the convergence of the state $x(t)$ for the initial values of

$\alpha_0 = .9$, and $.2$, respectively.

$a = 1, \gamma = 5, \beta = 1, \text{true } \alpha = .5, \text{initial } \alpha = .9$			
Iteration	α	$J(\alpha)$	$x(1)$
0	.9	78.9130445	-7.7572184
1	.8061489	11.8804011	-2.7175508
2	.7036760	1.6122215	-.7432785
3	.6046577	.1689001	.0158974
4	.5327346	.0090401	.2877378
5	.5034419	.0000834	.3602260
6	.5000500	.0000000	.3677829

Table 4.1

$a = 1, \gamma = 5, \beta = 1, \text{true } \alpha = .5, \text{initial } \alpha = .2$			
Iteration	α	$J(\alpha)$	$x(1)$
0	.2	.1205106	.637832
1	.8835656	55.8792983	-6.443195
2	.7876713	8.3050417	-2.201167
3	.6845883	1.0919149	-.542773
4	.5884893	.1055430	.090622
5	.5236488	.0045683	.311016
6	.5018820	.0000246	.363724
7	.5000214	.0000000	.367845

Table 4.2

Example 4.2. Here we set $a = 1, \beta = 1, \gamma = 4$, and $\alpha = .9$. The nonhomogeneous term is

$$f(t) = -1 - \gamma \frac{10e^{-t}}{\Gamma(.1)} t^{.1}.$$

In this case the true solution is e^{-t} .

This example contains a kernel that is more singular than that of Example 1. The results for initial values of $\alpha_0 = .999$ and $\alpha_0 = .8$ are given in Tables (4.3) and (4.4), respectively. For comparison, the true value of $x(1)$ is $x(1) = .3678794$. Note again it appears that a high initial guess of α is preferable

to a low one. Moreover, for an initial guess of $\alpha_0 = .75$ the algorithm updates the parameter to a value greater than 1 and the program stops. The convergence of the states is shown in Figures (4.3) and (4.4).

a = 1, $\gamma = 4$, $\beta = 1$, true $\alpha = .9$, initial $\alpha = .999$			
Iteration	α	J(α)	x(1)
0	.999	56.1016515	7.127846
1	.9400160	4.37204423	2.223429
2	.9075174	.1077531	.656523
3	.9002876	.0001462	.378490
4	.9000004	.0000000	.367895

Table 4.3

a = 1, $\gamma = 4$, $\beta = 1$, true $\alpha = .9$, initial $\alpha = .8$			
Iteration	α	J(α)	x(1)
0	.8	7.0279032	-1.895701
1	.9676194	17.3817062	4.096880
2	.9200856	.8805865	1.195951
3	.9019895	.0071213	.441968
4	.9000204	.0000007	.368632
5	.9000000	.0000000	.367879

Table 4.4

Example 4.3. This example has two features that are different than the previous examples. In Section 2 we assumed that $\beta > 0$, ensuring that the integral in equation (2.4) exists. This example illustrates that it may be possible to lift this restriction to include $\beta = 0$, which results in a true fractional derivative model. Also, for this example $x(t) = t^{1.5}$. Thus the true solution has an unbounded second derivative at $t = 0$. Because integration methods based on Simpson's rule converge slowly for functions that do not have four continuous derivatives, it was necessary to increase N to 200 for this example to gain accuracy.

Here we set the values $a = 1$, $\beta = 0$, $\gamma = 5$, and $\alpha = .5$. The function $f(t)$ is in this case

$$f(t) = \frac{3t^{.5}}{2} - \frac{2t^{2.5}}{5} - \Gamma(.5)\gamma\frac{3t^2}{8}.$$

The results of the quasilinearization algorithm are given in Table (4.5) for the initial value of $\alpha_0 = .9$. The fact that we could only obtain α to 3 correct digits is due to the inaccuracy of the integration scheme used. Figure (4.5) illustrates the convergence of the states to the true solution.

a = 1, $\gamma = 5$, $\beta = 0$, true $\alpha = .5$, initial $\alpha = .9$			
Iteration	α	J(α)	x(1)
0	.9	24.1608906	5.563632
1	.7641130	3.5153401	2.711820
2	.6246967	.3481690	1.531838
3	.5298167	.0127381	1.101062
4	.5019498	.0000392	1.005624
5	.5002170	.0000000	1.000050
6	.5002110	.0000000	1.000030

Table 4.5

References

1. R. L. Bagley and P. J. Torvik, *Fractional calculus -- a different approach to the analysis of viscoelastically damped structures*, AIAA Journal 21 (1983), 741-748.
2. H. T. Banks, R. H. Fabiano, and Y. Wang, *Estimation of Boltzmann damping coefficients in beam models*, LCDS/CCS Report No. 88-13, Brown University, Providence, 1988.
3. H. T. Banks and G. M. Groome, Jr., *Convergence theorems for parameter estimation by quasilinearization*, J. Math. Anal. Appl. 42 (1973), 91-109.
4. H. T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, 1989.
5. V. Barbu and S. I. Grossman, *Asymptotic behaviour of linear integrodifferential systems*, Trans. Amer. Math. Soc. 173 (1972), 277-288.
6. D. W. Brewer, *The differentiability with respect to a parameter of the solution of a linear abstract Cauchy problem*, SIAM J. Math. Anal. 13 (1982), 607-620.
7. D. W. Brewer, *A parameter dependence problem in functional differential equations*, Lecture Notes in Pure and Applied Mathematics, 81 (1982), Marcel Dekker, Inc., 187-195.
8. D. W. Brewer, J. A. Burns, and E. M. Cliff, *Parameter identification for an abstract Cauchy problem by quasilinearization*, ICASE Report No. 89-75, 1989.
9. J. A. Burns and R. H. Fabiano, *Modelling and approximation for a viscoelastic control problem*, Proc. Third International Conference on Control and Identification of Distributed Systems, Vorau, July 1986, 23-39.
10. J. A. Burns and T. L. Herdman, *Adjoint semigroup theory for a Volterra integrodifferential system*, Bull. Amer. Math. Soc. 81 (1975), 1099-1102.
11. J. A. Burns and T. L. Herdman, *Adjoint semigroup theory for a class of functional differential equations*, SIAM J. Math. Anal. 7 (1976), 729-745.
12. W. Desch and R. K. Miller, *Exponential stabilization of Volterra integral equations with singular kernels*, J. Integral Equations and Applications 1 (1988), 397-433.

13. Y. C. Fung, *Foundations of Solid Mechanics*, Prentice-Hall, Inc., 1965.
14. K. B. Hannsgen and R. L. Wheeler, *Time delays and boundary feedback stabilization in one-dimensional viscoelasticity*, Proc. Third International Conference on Control and Identification of Distributed Parameter Systems, Vorau, July 1986, 136-152.
15. W. J. Hrusa, J. A. Nohel, and M. Renardy, *Initial value problems in viscoelasticity*, Appl. Mech. Rev. 41 (1988), 371-378.
16. P. Linz, *Analytical and Numerical Methods for Volterra Equations*, SIAM Studies in Applied Mathematics, 1985.
17. M. Renardy, W. J. Hrusa, and J. A. Nohel, *Mathematical Problems in Viscoelasticity*, Pitman Monographs and Surveys in Pure and Applied Mathematics 35, Longman Scientific and Technical, Burnt Mill, 1987.
18. P. J. Torvik and R. L. Bagley, *On the appearance of the fractional derivative in the behaviour of real materials*, J. Appl. Mech. 51 (1984), 294-298.

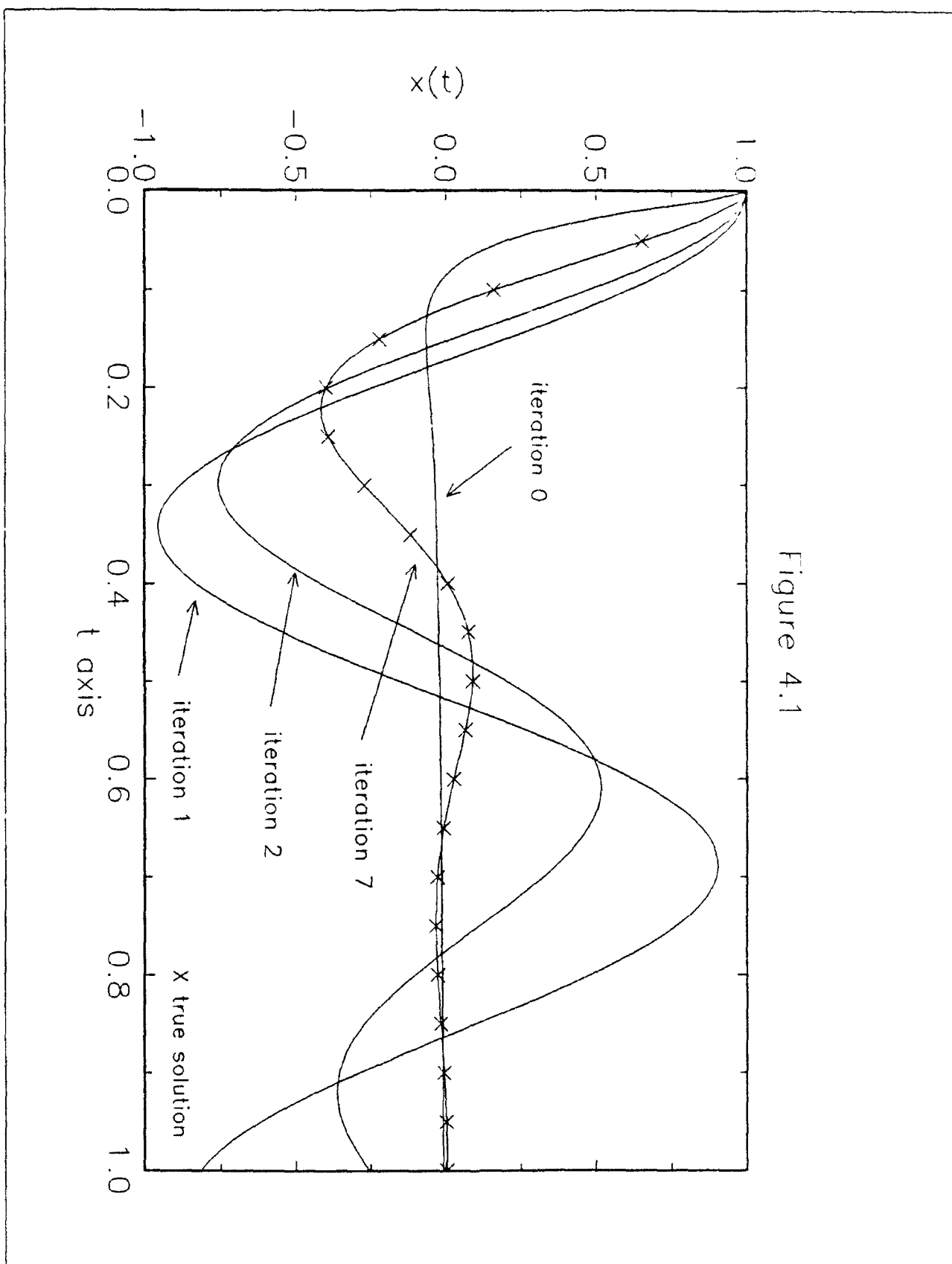


Figure 4.1

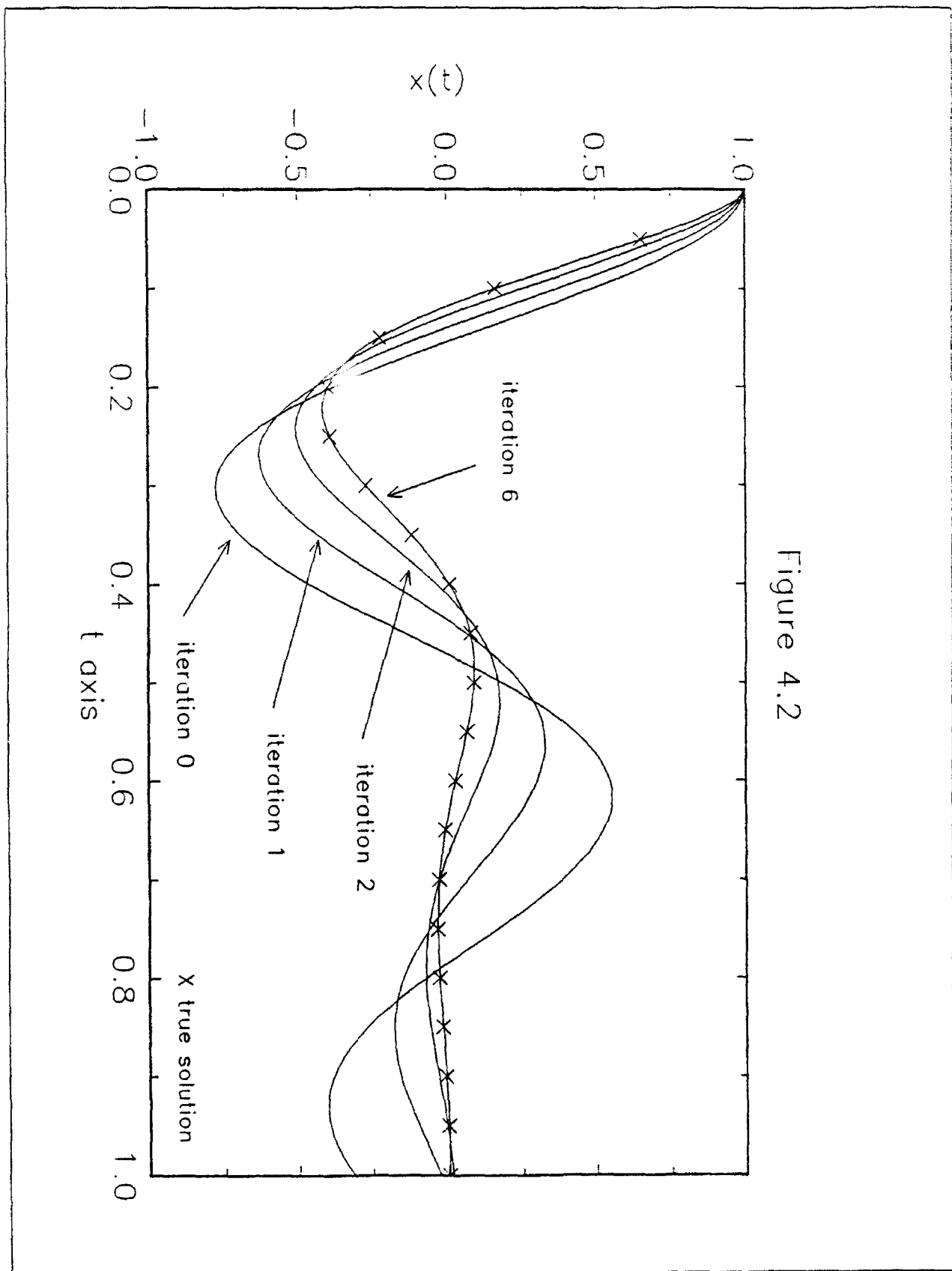


Figure 4.2

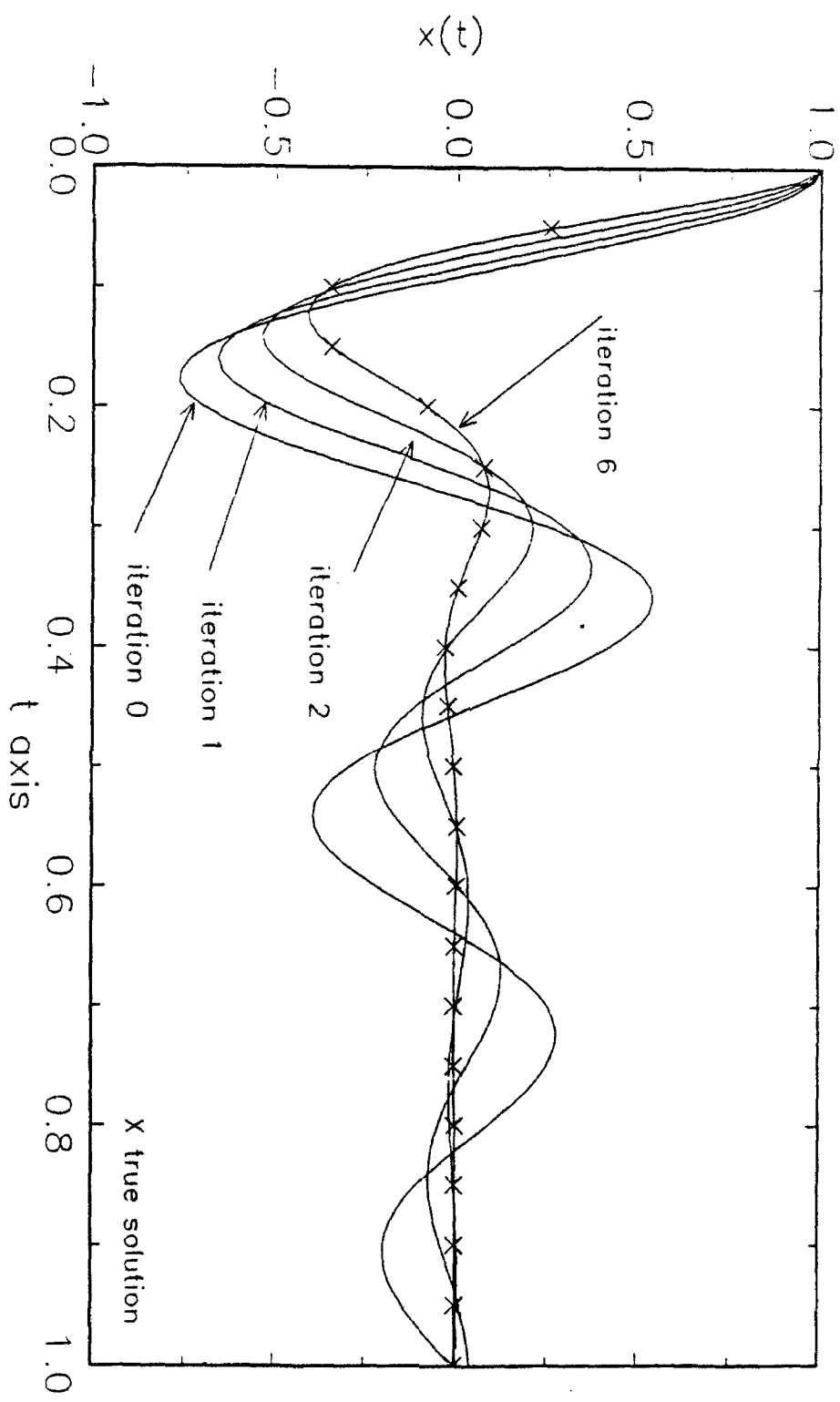
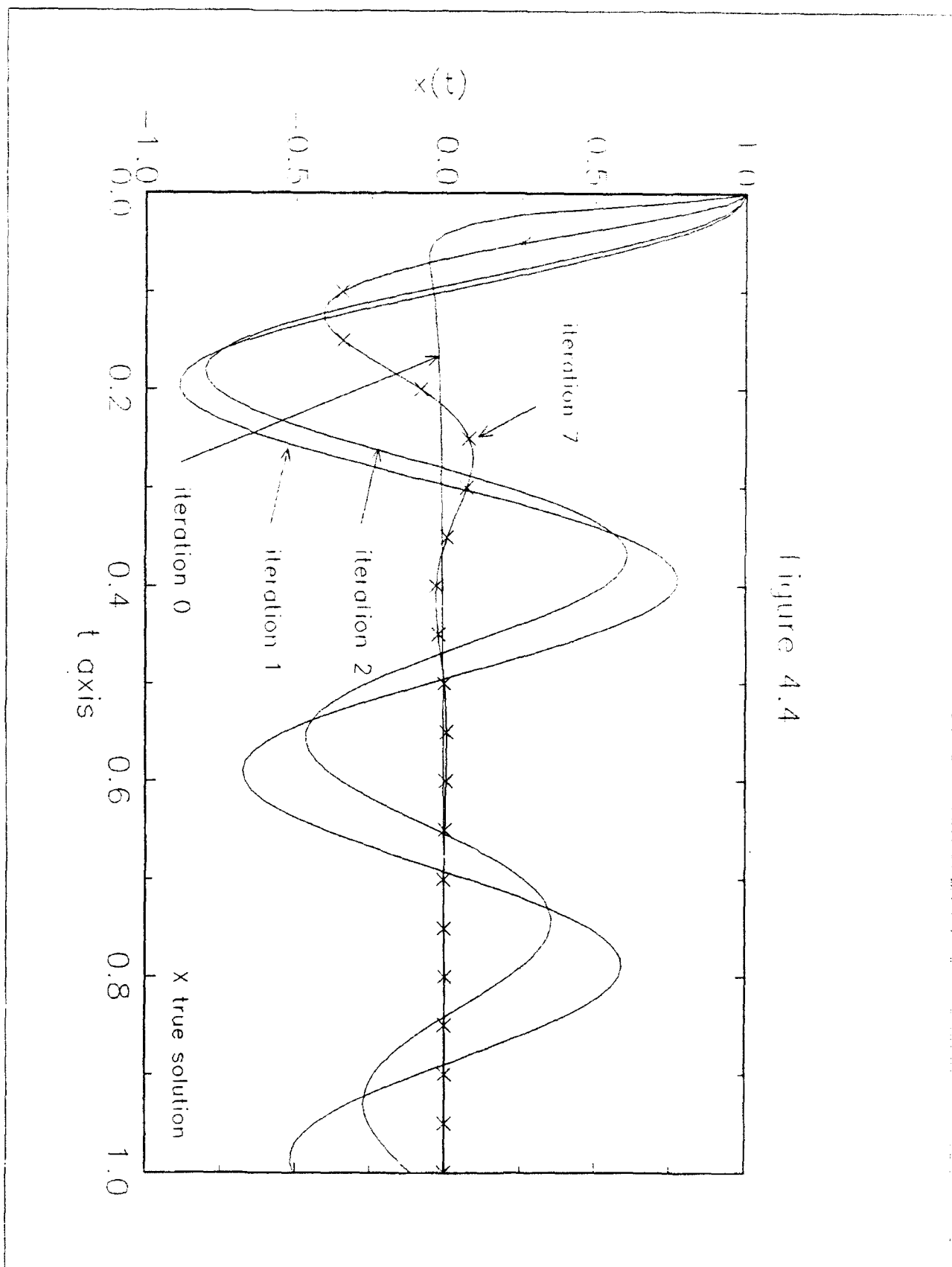


Figure 4.3



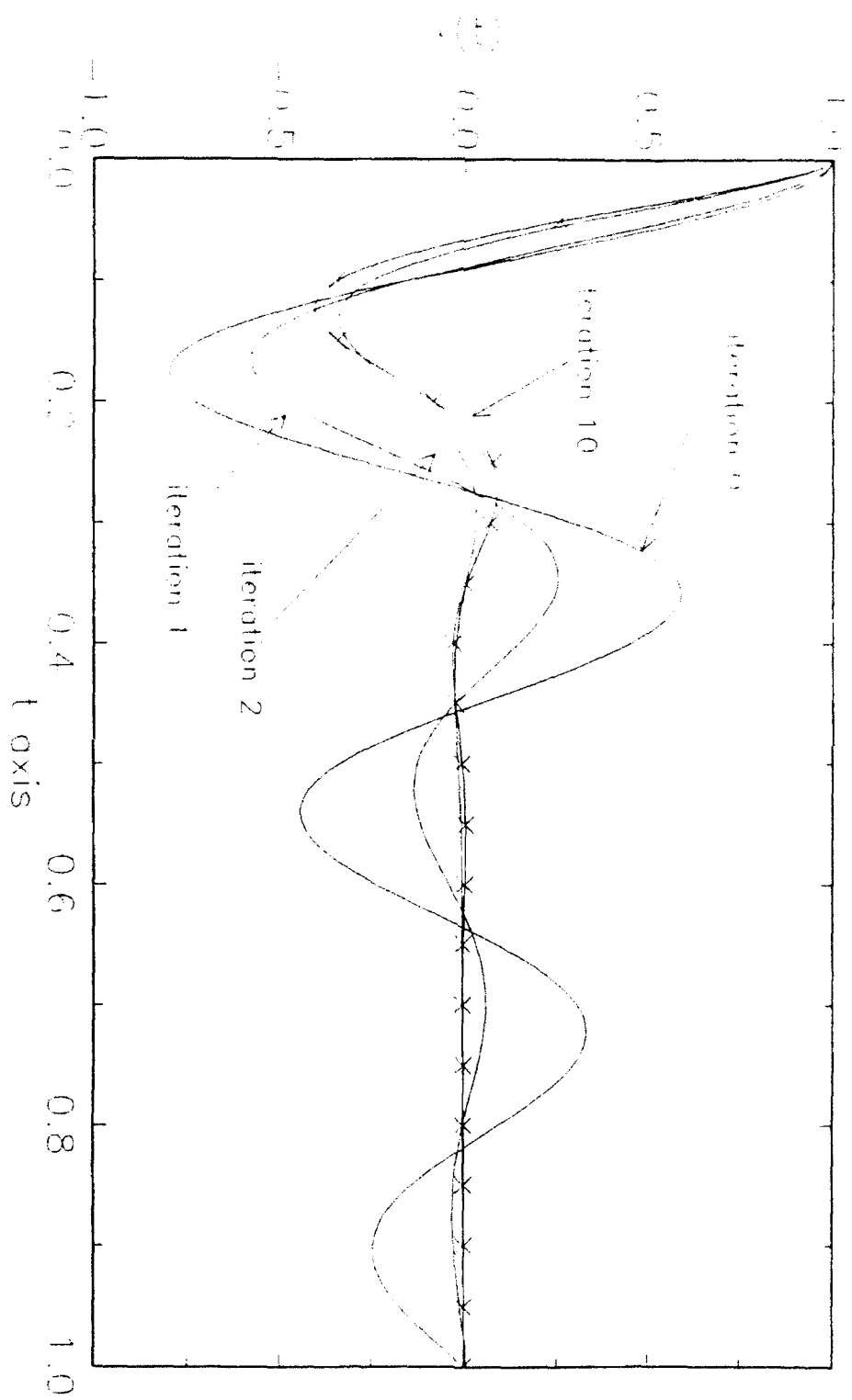


Figure 4.5

A DIRECT METHOD FOR PARAMETER ESTIMATION IN DISTRIBUTED SYSTEMS¹

Dennis W. Brewer

and

Robert K. Powers

Department of Mathematical Sciences
University of Arkansas
Fayetteville, Arkansas 72701

Abstract. We consider identification of parameters in a partial differential equation modeling the longitudinal vibrations of a viscoelastic beam. A semi-discrete approximation of this model gives rise to a Volterra integro-differential system with a weakly singular kernel. Such kernels arise in fractional derivative damping models of viscoelastic materials. A quasilinearization method is used to identify damping parameters in the system. We present the results of numerical experiments using a Galerkin approximation in space and an approximation in time which is adapted to weakly singular systems whose kernels are not continuous at the origin.

1. Introduction. In this paper we consider numerical methods for the identification of parameters in an idealized physical model for the longitudinal motions of a uniform bar fixed at both ends with Boltzmann type damping. The governing equation is ([10], [15])

$$(1.1) \quad \left\{ \begin{array}{l} \rho u_{tt}(x,t) = \frac{\partial}{\partial x} \left\{ E u_x(x,t) + \frac{\partial}{\partial t} \int_0^t g(t-s) u_x(x,s) ds \right\} \\ \quad + f(x,t), \quad 0 < x < 1, \quad t > 0, \end{array} \right.$$

with boundary conditions $u(0,t) = 0, \quad u(1,t) = 0,$

¹This research was supported by the Air Force Office of Scientific Research under grant AFOSR-89-0472.

and initial conditions $u(x,0) = d(x), \quad u_t(x,0) = v(x).$

Here, $u(x,t)$ represents the axial displacement of position x at time t , ρ is the density of the material, E a stiffness parameter, $f(x,t)$ a forcing function, and

$$g(s,q) = \frac{\gamma e^{-\beta s}}{\Gamma(1-\alpha)s^\alpha}, \quad s > 0,$$

represents a fractional derivative damping term modified to have exponential decay [13]. Here $\Gamma(\cdot)$ denotes the gamma function and $q = (\alpha, \beta) \in \mathbb{R}^2$ with $0 \leq \alpha < 1$ and $\beta > 0$. Such kernels arise in the study of fractional derivative models of viscoelastic structures. We refer the reader to [13], [19], [14], [16], and in particular to [18] and the extensive bibliography therein for a discussion of viscoelastic models as they relate to weakly singular kernels.

In this paper we consider the identification of α and β . Much of the groundwork for this paper appears in [9] and [8] where the identification of scalar systems is considered and in [6] and [7] where the theoretical framework used here is established. Torvik and Bagley [1], [19] have estimated the parameter α in the Laplace transform domain. Banks, et. al., [2], [4] have identified parameters corresponding to β and γ in a similar but different model assuming α is known.

A Galerkin approximation is used to approximate (1.1) by a Volterra equation with weakly singular kernel. In the example presented below, this semi-discrete model is solved using a product integration method based on Simpson's rule. Product integration methods using polynomial collocation are well suited for equations of the form (1.1) if the solution is analytic on the entire interval of interest. However, for the kernel of interest in this paper, the solution is not analytic at $t = 0$ if $f(t)$ is smooth [17].

If one applies a Galerkin scheme to equation (1.2), one obtains a system of second-order integro-differential equations. Integrating a scalar form of this system yields an equation of the form (See [9] for details.)

$$(1.2) \quad \begin{cases} \dot{x}(t) = a \int_0^t x(s)ds + \int_0^t g(t-s)x(s)ds + f(t), \\ x(0) = x_0. \end{cases}$$

A standard assumption in viscoelasticity [14] is that the material is in an unstrained state for time $t < 0$. This would correspond to $u(x,s) = 0$ for $s < 0$ in equation (1.1). Define $z(t) = \int_0^t x(s)ds$. It follows then that $x(s) = 0$ and $z(s) = 0$

for $s < 0$. If we define $w(t) = \text{col}(x(t), z(t))$, then $w(t)$ satisfies

$$(1.3) \quad \begin{cases} \dot{w}(t) = Mw(t) + \int_{-\infty}^t K(t-s)w(s)ds + F(t), \\ w(0) = \begin{pmatrix} x_0 \\ 0 \end{pmatrix}, \quad w(s) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad s < 0, \end{cases}$$

where $M = \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}$, $K(s) = \begin{pmatrix} g(s) & 0 \\ 0 & 0 \end{pmatrix}$, and $F(t) = \begin{pmatrix} f(t) \\ 0 \end{pmatrix}$.

2. The algorithm. In this section we state for reference the quasilinearization algorithm which is defined in an abstract semigroup formulation of equation (1.3). Details of this construction and related smoothness results may be found in [7] and [9]. We consider equation (1.3) in the form

$$(2.1) \quad \begin{cases} \dot{w}(t) = Mw(t) + \int_{-\infty}^0 K(-s, q)w(t+s)ds + F(t), \quad t > 0, \\ w(0) = \eta, \quad w(s) = \varphi(s), \quad s < 0, \end{cases}$$

where $\eta = \begin{pmatrix} x_0 \\ 0 \end{pmatrix} \in \mathbb{R}^2$, $M = \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}$, $F(t) = \begin{pmatrix} f(t) \\ 0 \end{pmatrix}$, and

$$K(s, q) = \frac{\gamma e^{-\beta s}}{\Gamma(1-\alpha)s^\alpha} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad s > 0.$$

We are interested in dependence on the parameter $q = (\alpha, \beta)$ where $\beta > 0$ and $0 \leq \alpha < 1$. By a solution of (2.1) we mean a function $w: (-\infty, \infty) \rightarrow \mathbb{R}^2$ such that w is absolutely continuous (A.C.) on $[0, \infty)$ and satisfies the integral equation a.e. on $[0, \infty)$, $w(0) = \eta$, and $w(s) = \varphi(s)$ a.e. on $(-\infty, 0]$.

Our semigroup formulation follows the construction in [5] as further developed in [11] and [12]. Define the product space $X = \mathbb{R}^2 \times L^1_f(-\infty, 0)$ with norm $\|(\eta, \varphi)\|_X = |\eta| + \|\varphi\|_{L^1_f(-\infty, 0)}$. Then it can be shown that the solution of (2.1) is given by the variation of constants formula

$$(w(t), w_t) = S(t, q)(\eta, \varphi) + Q(t, q)$$

where

$$(2.2) \quad Q(t, q) = \int_0^t S(t-s, q) (F(s), 0) ds.$$

and $S(t, q)$ is a C_0 -semigroup on X corresponding to the homogeneous version of (2.1).

We are now in a position to define a parameter estimation algorithm based on quasilinearization and state some local convergence results. For later adaptation we develop the algorithm for the case $q \in P \subset \mathbb{R}^n$ with canonical basis e_i , $i = 1, 2, \dots, n$. In Section 3 the algorithm is applied in the cases $n = 1$ and $n = 2$. The definitions and theorems stated here may also be found in [7], but are included here for immediate reference.

Let $y_0 = (\eta, \varphi) \in X$ and $q \in P$. Let C be a bounded linear mapping from X into a finite-dimensional space \mathbb{R}^ℓ , and define

$$y(t, q) = C[S(t, q)y_0 + Q(t, q)].$$

The parameter estimation algorithm is related to the following optimization problem.

Problem 2.1. Let $y_j \in \mathbb{R}^\ell$, $j = 1, 2, \dots, m$ be data values taken at times $t_j \in [0, T]$, $j = 1, 2, \dots, m$, respectively. For $q \in P$ define the quadratic cost function

$$J(q) = \sum_{j=1}^m |y(t_j, q) - y_j|^2.$$

Find $q^* \in P$ such that $J(q^*) \leq J(q)$ for all $q \in P$.

The quasilinearization algorithm method defines a recursive algorithm whose fixed point is a local solution of Problem 2.1. A more complete exposition is given in [3]. Given an initial guess $q_0 \in P$ define

$$q_{k+1} = \ell(q_k), \quad k = 0, 1, 2, \dots$$

where

$$\begin{aligned} \ell(q) &= q - [D(q)]^{-1}b(q) \\ D(q) &= \sum_{j=1}^m M^T(t_j, q)M(t_j, q) \\ b(q) &= \sum_{j=1}^m M^T(t_j, q)[y(t_j, q) - y_j] \end{aligned}$$

and the matrix $M(t, q)$ has its i^{th} column $M^i(t, q)$ given by

$$M^i(t, q) = CD_q[S(t, q)y_0 + Q(t, q)]e_i, \quad i = 1, 2, \dots, n.$$

The following theorems are typical of quasilinearization methods. Their proofs may be found in [7]. We obtain superlinear convergence when there is an exact fit to data (Theorem 2.1) and linear convergence in the presence of error (Theorem 2.2).

Theorem 2.1. Suppose the conditions of the previous section are satisfied. Moreover, assume $[D(q)]^{-1}$ exists, $f(q^*) = q^*$, and $J(q^*) = 0$. Then for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$|f(q) - f(q^*)| \leq \epsilon |q - q^*|$$

for $|q - q^*| \leq \delta$. In particular, there is a neighborhood \mathcal{U} of q^* such that $q_k \rightarrow q^*$ as $k \rightarrow \infty$ whenever $q_0 \in \mathcal{U}$.

The following theorem does not require an exact fit to data but does place some technical restrictions on the behaviour of the matrix $M(t, q)$ near q^* . Note that under the conditions of Theorem 2.2 there exists number $\bar{\delta} > 0$ such that for $0 < \delta < \bar{\delta}$ there exists a constant $K(\delta)$ such that

$$\sum_{j=1}^m |M^T(t_j, q) - M^T(t_j, q^*)| \leq K(\delta) |q - q^*|$$

for $|q - q^*| \leq \delta$. Let $K^* = \limsup_{\delta \downarrow 0} K(\delta)$ and define

$$\lambda^* = K^* |D(q^*)|^{-1} \max_j |w(t_j, q^*) - w_j|.$$

Theorem 2.2. Suppose the conditions of the previous section are satisfied. Moreover, assume $[D(q^*)]^{-1}$ exists and $f(q^*) = q^*$. Let λ^* be defined as above and assume $\lambda^* < 1$. Then there exists $\delta^* > 0$ such that

$$|f(q) - f(q^*)| \leq \lambda^* |q - q^*|$$

for $|q - q^*| \leq \delta^*$. In particular, $q_k \rightarrow q^*$ as $k \rightarrow \infty$ whenever $|q_0 - q^*| \leq \delta^*$.

3. Numerical results. In this section we present numerical results for parameter estimation in a semi-discrete approximation of the partial differential equation

$$\left\{ \begin{array}{l} \rho u_{tt}(x,t) = E u_{xx}(x,t) \\ \quad + \frac{\gamma}{\Gamma(1-\alpha)} \frac{\partial}{\partial t} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} u_{xx}(x,s) ds + f_1(x,t) \end{array} \right.$$

with boundary conditions $u(0,t) = 0$, $u(1,t) = 0$, and initial conditions $u(x,0) = u_0(x)$, $u_t(x,0) = u_1(x)$.

Integrating with respect to t one obtains

$$\left\{ \begin{array}{l} \rho u_t(x,t) = E \int_0^t u_{xx}(x,s) ds + \\ \quad \frac{\gamma}{\Gamma(1-\alpha)} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} u_{xx}(x,s) ds + f_2(x,t) \end{array} \right.$$

$$\text{where } f_2(x,t) = \int_0^t f_1(x,s) ds + \rho u_1(x).$$

We apply a Galerkin approximation in which the interval $[0,1]$ is divided into N equal parts and the homogeneous boundary conditions allow us to approximate the solution by a function

of the form $u(x,t) = \sum_{j=0}^N a_j(t) \phi_j(x)$ where ϕ_j is a cubic spline basis element. Substituting in the equation, taking an inner product with a basis element ϕ_k , and integrating by parts yields a Volterra equation in t of the form

$$(3.1) \quad \left\{ \begin{array}{l} \rho A^N \dot{v}(t) = -E D^N \int_0^t v(s) ds - \\ \quad \frac{\gamma}{\Gamma(1-\alpha)} D^N \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} v(s) ds + f_3(t) \end{array} \right.$$

where A^N and D^N are $(N+1) \times (N+1)$ matrices depending on inner products of ϕ_j and ϕ_j' .

The quasilinearization algorithm requires that we solve (3.1) along with its sensitivity equations which have the form

$$(3.2) \quad \left\{ \begin{aligned} \dot{x}_\alpha(t) &= a \int_0^t x_\alpha(s) ds + \frac{\gamma}{\Gamma(1-\alpha)} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} x_\alpha(s) ds \\ &\quad + \frac{\gamma \Gamma'(1-\alpha)}{\Gamma(1-\alpha)^2} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} x(s) ds \\ &\quad - \frac{\gamma}{\Gamma(1-\alpha)} \int_0^t \frac{\ln(t-s) e^{-\beta(t-s)}}{(t-s)^\alpha} x(s) ds, \\ x_\alpha(0) &= 0, \end{aligned} \right.$$

$$(3.3) \quad \left\{ \begin{aligned} \dot{x}_\beta(t) &= a \int_0^t x_\beta(s) ds + \frac{\gamma}{\Gamma(1-\alpha)} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} x_\beta(s) ds \\ &\quad - \frac{\gamma}{\Gamma(1-\alpha)} \int_0^t \frac{e^{-\beta(t-s)}(t-s)}{(t-s)^\alpha} x(s) ds \\ x_\beta(0) &= 0. \end{aligned} \right.$$

where $x_\alpha = \frac{\partial x}{\partial \alpha}$ and $x_\beta = \frac{\partial x}{\partial \beta}$. The zero initial condition reflects the fact that $x(0) = x_0$ is independent of α and β . It is important to note that (3.2) and (3.3) also have weakly singular kernels.

The implementation of the identification scheme begins with an initial guess for $q = (\alpha, \beta)$. Equations (3.1)-(3.3) are integrated using this initial value, then $x(t)$, $x_\alpha(t)$, and $x_\beta(t)$ are used to give an updated estimate of the parameter using the quasilinearization algorithm.

As a specific example consider the partial differential equation

$$\left\{ \begin{aligned} \rho u_{tt}(x, t) &= E u_{xx}(x, t) \\ &\quad + \frac{\gamma}{\Gamma(1-\alpha)} \frac{\partial}{\partial t} \int_0^t \frac{e^{-\beta(t-s)}}{(t-s)^\alpha} u_{xx}(x, s) ds + f(x, t) \end{aligned} \right.$$

with boundary conditions $u(0, t) = 0$, $u(1, t) = 0$, and initial conditions $u(x, 0) = 0$, $u_t(x, 0) = 4\pi \sin(2\pi x)$, where $\rho = 1$,

$E = 1$, $\gamma = 1$, $\alpha = 1/2$, and $\beta = 0$. The exact solution is $u(x, t) = \sin(2\pi x) \sin(4\pi t)$ with the forcing function chosen as

$$f(x,t) = \pi \sin(2\pi x) + 3\pi \sin(2\pi x) \cos(4\pi t) + 4\pi^2 \sqrt{t} \sin(4\pi x) \sum_{k=0}^{\infty} \frac{(-1)^k (b \cdot t)^{2k+1}}{\Gamma(2k + 5/2)}.$$

The numerical scheme uses $N = 8$ and $h = .02$. The parameter estimator uses 70 data points taken from the exact form of $u_t(x,t)$. The results of an application of the quasilinearization algorithm to the estimation of the parameter α are given in the table below.

Iteration	α	$J(\alpha)$
0	.9	1703.2400606
1	0.0783169	10811.4763257
2	0.2310107	4486.6256214
3	0.3658041	1012.4621853
4	0.4729954	32.7112404
5	0.4990789	0.0523379
6	0.5002075	0.0001728
7	0.5002095	0.0001726
8	0.5002095	0.0001726

The table below shows the results of the algorithm for identifying both the parameter α and β .

Iteration	α	β	$J(\alpha, \beta)$
0	0.2000000	0.5000000	5076.0182126
1	0.2986892	6.5002076	1394.6849127
2	0.5281903	7.7538270	273.4823741
3	0.4674188	-4.2094682	469.6983381
4	0.5135026	-3.4408819	70.7243517
5	0.5203233	-2.3067455	26.1540142
6	0.5029341	-0.1314789	0.2291736
7	0.5001956	0.0032037	0.0001997
8	0.5002027	0.0009508	0.0001674

References

1. R. L. Bagley and P. J. Torvik, *Fractional calculus -- a different approach to the analysis of viscoelastically damped structures*, AIAA Journal 21 (1983), 741-748.
2. H. T. Banks, R. H. Fabiano, and Y. Wang, *Estimation of Boltzmann damping coefficients in beam models*, LCDS/CCS Report No. 88-13, Brown University, Providence, 1988.

3. H. T. Banks and G. M. Groome, Jr., *Convergence theorems for parameter estimation by quasilinearization*, J. Math. Anal. Appl. 42 (1973), 91-109.
4. H. T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, 1989.
5. V. Barbu and S. I. Grossman, *Asymptotic behaviour of linear integrodifferential systems*, Trans. Amer. Math. Soc. 173 (1972), 277-288.
6. D. W. Brewer, *The differentiability with respect to a parameter of the solution of a linear abstract Cauchy problem*, SIAM J. Math. Anal. 13 (1982), 607-620.
7. D. W. Brewer, J. A. Burns, and E. M. Cliff, *Parameter identification for an abstract Cauchy problem by quasilinearization*, Q. Appl. Math., to appear.
8. D. W. Brewer and R. K. Powers, *Parameter estimation for a Volterra integro-differential equation*, Appl. Num. Math. 9 (1992), 307-320.
9. D. W. Brewer and R. K. Powers, *Parameter identification in a Volterra equation with weakly singular kernel*, J. Int. Eqns. Appl. 2 (1990), 353-373.
10. J. A. Burns and R. H. Fabiano, *Modelling and approximation for a viscoelastic control problem*, Proc. Third International Conference on Control and Identification of Distributed Systems, Vorau, July 1986, 23-39.
11. J. A. Burns and T. L. Herdman, *Adjoint semigroup theory for a Volterra integrodifferential system*, Bull. Amer. Math. Soc. 81 (1975), 1099-1102.
12. J. A. Burns and T. L. Herdman, *Adjoint semigroup theory for a class of functional differential equations*, SIAM J. Math. Anal. 7 (1976), 729-745.
13. W. Desch and R. K. Miller, *Exponential stabilization of Volterra integral equations with singular kernels*, J. Integral Equations and Applications 1 (1988), 397-433.
14. Y. C. Fung, *Foundations of Solid Mechanics*, Prentice-Hall, Inc., 1965.

15. K. B. Hannsgen and R. L. Wheeler, *Time delays and boundary feedback stabilization in one-dimensional viscoelasticity*, Proc. Third International Conference on Control and Identification of Distributed Parameter Systems, Vorau, July 1986, 136-152.
16. W. J. Hrusa, J. A. Nohel, and M. Renardy, *Initial value problems in viscoelasticity*, Appl. Mech. Rev. **41** (1988), 371-378.
17. R. K. Miller and A. Feldstein, *Smoothness of solutions of Volterra integral equations with weakly singular kernels*, SIAM J. Math. Anal. **2** (1971), 242-258.
18. M. Renardy, W. J. Hrusa, and J. A. Nohel, *Mathematical Problems in Viscoelasticity*, Pitman Monographs and Surveys in Pure and Applied Mathematics 35, Longman Scientific and Technical, Burnt Mill, 1987.
19. P. J. Torvik and R. L. Bagley, *On the appearance of the fractional derivative in the behaviour of real materials*, J. Appl. Mech. **51** (1984), 294-298.